Journal of Applied Statistics
Vol. 00, No. 00, Month 200x, 1–17

RESEARCH ARTICLE

A flexible parametric survival model which allows a bathtub shaped hazard rate function.

William J. Reed[a] *

[a] *Department of Mathematics and Statistics, University of Victoria, PO Box 3060, Victoria, B.C., Canada V8W 3R4*;

(*July, 2009*)

A new parametric (3-parameter) survival distribution, the *lognormal-power function* distribution, with flexible behaviour is introduced. Its hazard rate function can be either unimodal, monotonically decreasing or can exhibit a bathtub shape. Special cases include the lognormal distribution and the power function distribution, with finite support. Regions of parameter space where the various forms of the hazard-rate function prevail are established analytically. The distribution lends itself readily to accelerated life (AL) regression modelling. Applications to five datasets taken from the literature are given. Also it is shown how the distribution can behave like a Weibull distribution (with negative aging) for certain parameter values.

Keywords: lognormal-power function distribution; bathtub hazard; AL regression; parametric hazard rate function; survival analysis.

1. Introduction.

The parametric survival distributions most commonly used in regression modelling (*e.g. Weibull, lognormal, log-logistic etc.*) have unimodal or monotone hazard rate functions and are incapable of modelling a bathtub shaped hazard, in spite of the fact that such bathtub hazards are quite common, especially in studies of animal survival from birth and of the failure of certain types of equipment.

Notable exceptions to this are the three-parameter *generalized Weibull* (GW) distribution [10] which not only can be used in accelerated life (AL) regression modelling, but also permits fully parametric proportional hazards (PH) regression modelling; the three-parameter *generalized gamma* (GG) distribution (see *e.g.* [2]) and the four-parameter *generalized F* (GF) distribution (see *e.g.* [3]). One difficulty with the GW distribution is the fact that in the case of a bathtub hazard it has finite support and the maximum likelihood (ML) estimate of the threshold parameter is an order statistic. In consequence the standard ML second-order asymptotics do not hold. ML estimation for the GG and GF distributions require numerical optimization, and problems involving the convergence of numerical routines have been reported, although the problem is greatly ameliorated by using the re-parameterization suggested by Prentice [12].

In this paper a new parametric distribution, with three parameters, is introduced for which the hazard rate can exhibit a bathtub shape. In addition it can be unimodal or monotonically decreasing. The distribution is that of the product

*Corresponding author. Email: reed@math.uvic.ca

ISSN: 0266-4763 print/ISSN 1360-0532 online

DOI: 10.1080/0266476YYxxxxxxxx
http://www.informaworld.com

of independent random variables, one following a lognormal distribution, and the other a power-function distribution with finite support. It turns out to be a special case (with the right-hand tail parameter set to infinity) of the *double-Pareto lognormal* distribution introduced in [13]. The new distribution can be thought of as arising from a (frailty-type) mixture model where each individual in the population has a lognormal, LN(μ, σ^2), survival distribution, with the individual medians (e^μ) across the population following a power-law distribution.

The new distribution provides an alternative to the GG, GF and GW distributions for parametric modelling of data exhibiting a bathtub hazard. Unlike the bathtub form of the GW, the new distribution does not suffer from the non-regularity in ML estimation (except in one special case when one parameter is zero), and experience has shown that numerical maximization of the likelihood function works well.

In the following section the distribution is defined and some basic properties derived. Section 3 summarizes the possible shapes of the hazard-rate function (proofs are in an Appendix). Regression formulations and estimation are discussed in the short Sections 4 and 5, while Section 6 presents examples using a variety of survival data taken from the literature. Comparisons of the fit with that of the 3-parameter GW and GG distributions are given for examples involving bathtub hazards.

2. The model.

We consider a model for survival time T with

$$Y = \log T = \mu + \sigma Z - \frac{1}{\beta} E \tag{1}$$

where σ and β are non-negative parameters, μ is a real parameter and Z and E are independent and follow respectively standard normal and standard exponential distributions *i.e.* $Z \sim \mathrm{N}(0,1)$ and $\mathrm{E} \sim \mathrm{EXPON}(1)$. The distribution of Y is in fact a special case of a *normal-Laplace* distribution [13] with the right-tail parameter α set equal to infinity *i.e.* $Y \sim \mathrm{NL}(\mu, \sigma^2, \infty, \beta)$. Properties of the normal-Laplace distribution can be invoked to establish the survivor function and probability density function (pdf) of Y as

$$S_Y(y) = \phi\left(\frac{y-\mu}{\sigma}\right)\left[R\left(\frac{y-\mu}{\sigma}\right) - R\left(\beta\sigma + \frac{y-\mu}{\sigma}\right)\right] \tag{2}$$

and

$$f_Y(y) = \beta\phi\left(\frac{y-\mu}{\sigma}\right) R\left(\beta\sigma + \frac{y-\mu}{\sigma}\right) \tag{3}$$

where R is *Mills' ratio* of the complementary cumulative distribution function (cdf) to the pdf of a standard normal distribution:

$$R(z) = \frac{\Phi^c(z)}{\phi(z)}.$$

Alternatively, avoiding the use of Mills' ratio,

$$S_Y(y) = \Phi^c\left(\frac{y-\mu}{\sigma}\right) - \exp\{\beta(y-\mu) + \beta^2\sigma^2/2\}\ \Phi^c\left(\beta\sigma + \frac{y-\mu}{\sigma}\right) \tag{4}$$

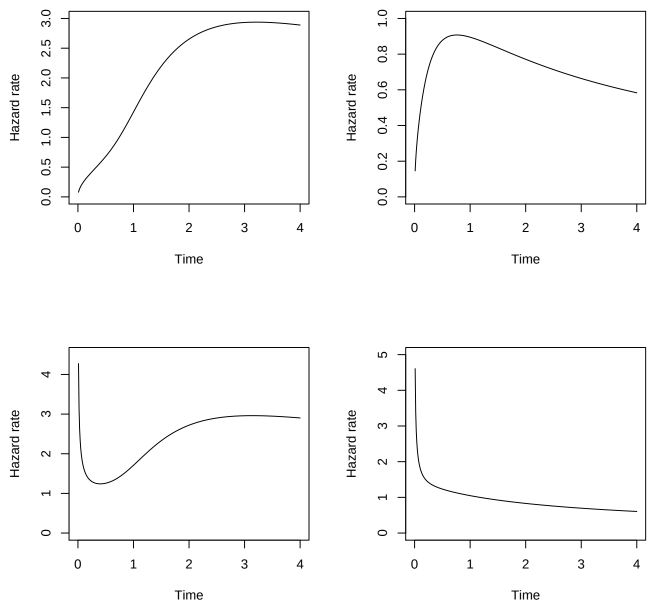

Figure 1. Some possible shapes of the hazard rate function for the lognormal-power function distribution. For all plots $\mu = 0.5$, while $\beta = 1.5$ (top row) with $\beta = 0.5$ (bottom row) and $\sigma = 0.3$ (left hand panels) with $\sigma = 0.8$ (right-hand panels).

$$f_Y(y) = \beta \exp\{\beta(y-\mu) + \beta^2\sigma^2/2\}\ \Phi^c\left(\beta\sigma + \frac{y-\mu}{\sigma}\right) \tag{5}$$

The corresponding survivor function and density of T are

$$S_T(t) = S_Y(\log t) \quad \text{and} \quad f_T(t) = \frac{1}{t} f_Y(\log t) \tag{6}$$

The hazard-rate function for T is

$$h(t) = \frac{\beta R(\beta\sigma + (\log t - \mu)/\sigma)}{t\left[R((\log t - \mu)/\sigma) - R(\beta\sigma + (\log t - \mu)/\sigma)\right]}. \tag{7}$$

The shape of this hazard-rate function is quite flexible (see Figure 1) and can be: monotone decreasing; unimodal; or of a form which has two turning points, at first decreasing to a local minimum then increasing to a local maximum before finally decreasing. In this last case, the hazard-rate exhibits a bathtub shape over the early part of the range.

The mean and variance of the survival time T are:

$$\mathrm{E}(T) = \frac{\beta}{\beta+1} e^{\mu+\sigma^2/2}; \qquad \mathrm{var}(T) = \beta e^{2\mu+\sigma^2}\left[\frac{e^{\sigma^2}}{\beta+2} - \frac{\beta}{(\beta+1)^2}\right]$$

Two special cases of the distribution governed by (1) warrant attention. The first arises in the limit as $\beta \to \infty$. From (1) it is seen that in this case the survival distribution is simply the *lognormal*. The second arises when $\sigma = 0$. In this case T

has a *power-function* distribution with finite support on $(0, e^{\mu})$ and pdf

$$f(t) = \beta e^{-\beta\mu} t^{\beta-1}. \tag{8}$$

thereon. The hazard-rate function is

$$h(t) = \frac{\beta t^{\beta-1}}{e^{\beta\mu} - t^{\beta}} \qquad \text{for } 0 < t < e^{\mu} \tag{9}$$

This hazard rate has a vertical asymptote at $t = e^{\mu}$. Furthermore if $\beta > 1$ it is increasing from 0 over $(0, e^{\mu})$; while if $\beta < 1$ it is U-shaped on $(0, e^{\mu})$, with vertical asymptotes at 0 and e^{μ} and a minimum at $(1-\beta)^{1/\beta} e^{\mu}$; in the intermediate case $(\beta = 1)$ it increases from $e^{-\mu}$.

We shall refer the distribution of T defined by (1) as a *lognormal-power function* (lNpf) distribution, because the distribution of such a T can be represented as a product of independent rvs with respectively lognormal and power-function distributions. Likewise we shall refer to the hazard-rate (9) as a *lognormal-power function* hazard rate.

The upper tail of this distribution behaves like the lognormal distribution, but there is much greater flexibility in the left-hand part of the distribution $(t < e^{\mu})$ including the possibility of a bathtub shaped hazard. This is discussed further in the following Section.

The lognormal-power function distribution can be thought of as arising from a (frailty-type) mixture model where each individual in the population has a lognormal, $\mathrm{LN}(m, \sigma^2)$ survival distribution, with the parameter m being distributed across the population as $\mu - E/\beta$, *i.e.* with density

$$f(m) = \beta e^{-\beta\mu} e^{\beta m} \qquad \text{for } m < \mu.$$

For such a model the individual median survival times, $\exp(m) = t_{0.5}$ say, would be distributed across the population with density

$$f(t_{0.5}) = \beta e^{-\beta\mu}\, t_{0.5}^{\beta-1} \qquad \text{for } 0 < t_{0.5} < e^{\mu},$$

a power-function distribution. If $\beta < 1$ this is decreasing (more individuals with shorter median lifetimes); if $\beta > 1$ it is increasing (more individuals with longer median lifetimes); and if $\beta = 1$ it is constant (a uniform distribution of median lifetimes). It will be seen in the next Section that the shape of the hazard-rate function (7) has different forms depending on how β compares with 1.

3. Some properties of the lognormal-power function hazard rate function.

Glaser [5] and Marshall & Olkin [8] provide sufficient conditions for a hazard-rate function to be monotone increasing or decreasing, unimodal or of bathtub shape. Unfortunately these conditions are not easy to apply for the lNpf model. In the Appendix other methods are used to examine the behaviour of the hazard-rate function. There it is shown that β is a critical parameter. If $\beta > 1$ (and $\sigma > 0$) the hazard-rate is qualitatively shaped like that of the lognormal distribution *i.e.* it is unimodal, with $h(0) = 0$ and $h(t) \to 0$ as $t \to \infty$. On the other hand if $\beta < 1$, the hazard-rate function has a vertical asymptote at zero and, depending on the value of σ is either (a) monotonically decreasing to a limiting value of zero (as $t \to \infty$),

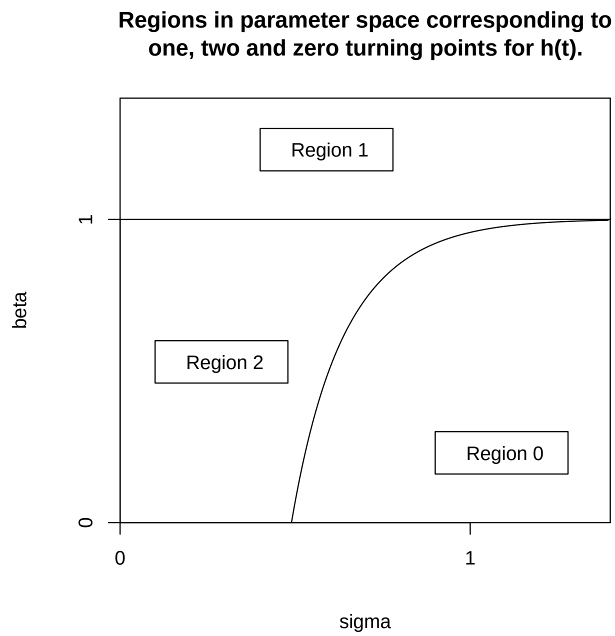

Figure 2. How parameter values determine the shape of the hazard rate function. In Region 1 ($\beta > 1$), the hazard rate function is unimodal; in Region 2 ($\beta < 1$ and $\sigma < \sigma^\star(\beta)$) it has two turning points, and can exhibit a bathtub shape for smaller vales of t while in Region 0 ($\beta < 1$ and $\sigma > \sigma^\star(\beta)$) the hazard rate is monotonically decreasing (0 turning points).

or (b) at first decreases to a local minimum, then increases to a local maximum and subsequently decreases to a limiting value of zero (as $t \to \infty$). In this latter case, for values of t less than the local maximum, the hazard-rate function has a bathtub shape. The latter case (b) occurs for σ suitably small ($\sigma < \sigma^\star(\beta)$) and the former case (a) for $\sigma > \sigma^\star(\beta)$. The partition of the parameter space into three regions corresponding to different forms of the hazard-rate function is shown in Figure 2. The examples presented in Figure 1, correspond to four points on a rectangular grid in β-σ parameter space.

In the Appendix it is shown how the boundary ($\beta = \sigma^\star(\beta)$) separating Regions 0 and 2 (corresponding to forms (a) and (b) above) can be computed numerically. When $\sigma = 0$ the hazard rate is either increasing ($\beta \geq 1$) or bathtub shaped ($\beta < 1$) as mentioned in the previous section.

The lNpf distribution can mimic the Weibull distribution for certain parameter values. Figure 3 shows an example of this. Plotted are the hazard-rate function, $h(t)$, and a doubly-logarithmic plot of the cumulative hazard $H(t) = -\log(S(t))$ against t, for parameter values $\mu = 15, \sigma = .7$ and $\beta = .2$. For a Weibull hazard, $\log H(t)$ is linear in $\log(t)$. The linearity of the plot in the right-hand panel of Figure 3 demonstrates the closeness of the lNpf hazard-rate to that of a Weibull distribution. In the Appendix it is shown analytically how for large μ and small $\beta\sigma$ the lNpf distribution will behave like a Weibull distribution for $t << e^\mu$.

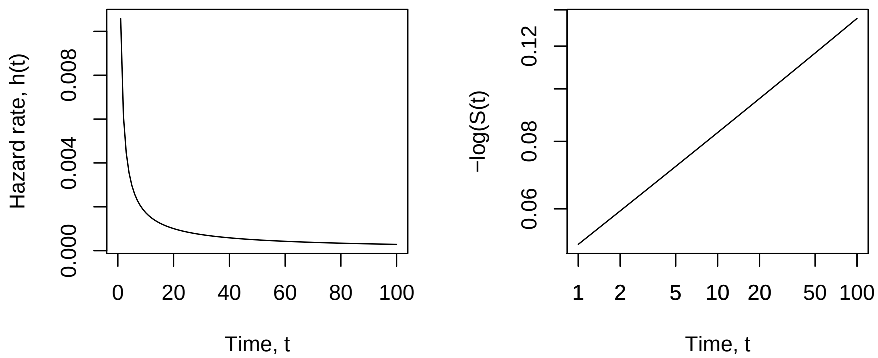

Figure 3. An illustration of how the lNpf distribution can behave like a Weibull distribution over a limited range. Plotted are the hazard rate function (l. hand panel) and the cumulative hazard, with both axes logarithmic (r. hand panel) for the lNpf distribution with $\mu = 15$, $\sigma = 0.7$ and $\beta = 0.2$. The fact that the logarithmic plot of the cumulative hazard is virtually linear over the range shown, indicates how well it mimics a Weibull distribution over this range.

4. Regression models.

The construction of an *accelerated life* (AL) regression model is straightforward from (1). With covariates $\mathbf{X}$, one can write the AL model as

$$Y = \log T = \boldsymbol{\gamma}^T \mathbf{X} + \sigma Z - \frac{1}{\beta} E \qquad (10)$$

where $\boldsymbol{\gamma}^T$ is a vector of regression coefficients, including an intercept term. Again the mixture interpretation of Sec.2 can be applied *i.e.* one can consider each unit following a lognormal survival time with the mean parameter (of the normal log(survival time)) depending on covariates and a random effect $-E/\beta$.

The class of lNpf models is not closed under the *proportional hazards* assumption. However an extended four-parameter class of hazard models with

$$h(t) = \theta h_T(t) \qquad S(t) = [S_T(t)]^\theta, \qquad (11)$$

where h_T and S_T are the lNpf hazard rate (7) and survivor (6) functions, is closed in this way, so that a fully parametric PH regression specification (in which the covariates influence θ) can be obtained.

5. Estimation.

Maximum likelihood (ML) estimation can be conducted in the usual way for survival time data. If there are no covariates and exact failure times $t_1, t_2, \ldots, t_d$ along with $\tau_1, \tau_2, \ldots, , \tau_{n-d}$ (right) censoring times are observed, the log-likelihood function can be computed as

$$\ell(\mu, \sigma, \beta) = \sum_{i=1}^{d} \log f_Y(\log t_i) + \sum_{j=1}^{n-d} \log S_Y(\log \tau_j) \qquad (12)$$

where f_Y and S_Y are as given in (3) and (2). If there are covariates $\mathbf{X}$ present, the log-likelihood (in $\boldsymbol{\gamma}^T$, σ and β) for the AL regression model is of the same form, but with μ replaced $\boldsymbol{\gamma}^T\mathbf{X}$.

For grouped survival data, with f_i failing between times $t_{(i)}$ and $t_{(i+1)}$, the log-likelihood is

$$\ell(\mu, \sigma, \beta) = \sum_{i=0}^{m-1} f_i \log[S_Y(\log t_{(i)}) - S_Y(\log(t_{(i+1)})] \tag{13}$$

where $t_0 = 0$ and $t_{(m)} = \infty$.

In computing the likelihood it is probably easiest to use the forms (4) and (5) avoiding Mills ratio and using exp(pnorm(x, lower=FALSE, log=TRUE) in R or S-Plus to compute $\Phi^c(x)$. However very good approximations can be obtained for Mills ratio using the expansion in terms of Tchebycheff-Hermite polynomials presented by Ruben (1962).

The log-likelihood must be maximized numerically. As starting values one can first fit a log-normal distribution and use the estimates obtained along with an arbitary positive value for β *e.g.* 1. For the examples in the following section the Nelder-Mead simplex algorithm was used with the R routine optim. In fitting the model to many datasets the author has not experienced any difficulty in obtaining convergence (except in cases when the optimum is on the boundary of parameter space) within a minute or two. In the two examples below (6.3 and 6.4) where the optimum is on the boundary of parameter space, the maximization routine failed to converge, stopping near to the boundary. Refitting the appropriate reduced model (power function or lognormal) then yielded a maximized log-likelihood equal to (or very slightly higher) than that obtained when the optimization routine stopped.

6. Examples.

In this section the lNpf distribution is applied to five datasets taken from the literature. All but Example 4 involve bathtub shaped hazards. The examples include both grouped and ungrouped data and some involve covariates with examples of both AL and PH regression. In two of the examples (Examples 3 and 4) the best fitting lNpf distribution turns out to be respectively a power function distribution and a lognormal distribution.

6.1 *Electrical appliances.*

Lawless [7] presents data on the number of cycles until failure for sixty electrical appliances in a life test. He fitted, in a somewhat informal way, a two-component mixture of Weibull distributions. Fitting an lNpf distribution yields ML estimates (with asymptotic standard errors): $\hat{\mu} = 8.539$ (0.137), $\hat{\sigma} = 0.340$ (0.077) and $\hat{\beta} = 0.679$ (0.230). The asymptotic correlations between $\hat{\mu}$ and both $\hat{\sigma}$ and $\hat{\beta}$ are negative (-0.573 and -0.574, respectively), while that between $\hat{\sigma}$ and $\hat{\beta}$ is positive (0.341). The maximized log-likelihood had a value of 519.205. A glance at Figure 2 reveals that the point in parameter space $(\hat{\sigma}, \hat{\beta})$ falls in Region 2, so the fitted hazard will have two turning points. This fitted hazard is plotted in Figure 4 (right) along with the fitted survivor function and Kaplan-Meier estimate (left). Lawless computes expected frequencies for the fitted Weibull mixture, after grouping the data into 9 cells. Using the same grouping, the fitted frequencies computed for the lNpf model turn out to be closer to the observed frequencies than those of the

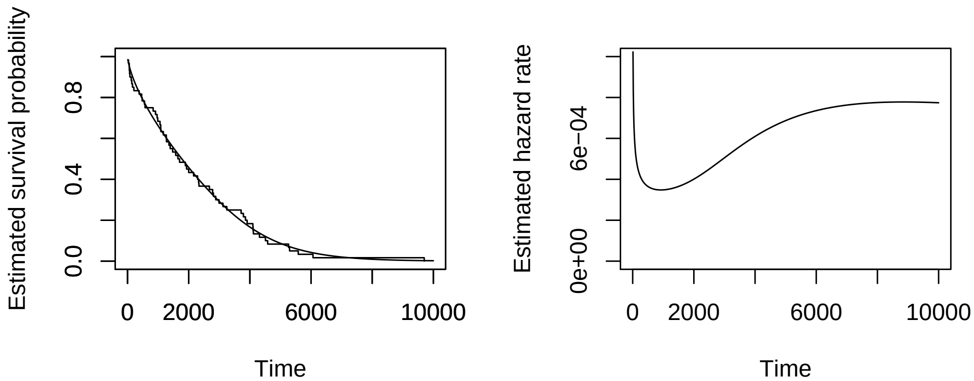

Figure 4. *Example 1. Electrical appliances.* The left-hand panel shows the fitted (MLE) lNpf survivor function and the non-parametric Kaplan-Meier estimate of the survivor function. The right-hand panel shows the fitted (MLE) hazard-rate function. Note the bathtub shape.

Weibull mixture yielding a Pearson χ^2 statistic of 2.87 (5 df) compared with 6.84 (3 df) for the Weibull mixture. Using the former value in a χ^2 goodness of fit test yields a P-value of 0.72, confirming the apparent good fit seen in Figure 4.

Fitting the (3-parameter) generalized Weibull (GW) and (3-parameter) generalized gamma (GG) distributions by maximum likelihood also yielded bathtub shaped fitted hazard-rate functions. The resulting values of the maximized log-likelihood were 519.15 and 519.59 respectively. Thus in terms of the Akaike Information Criterion (AIC) the lNpf, GW and GG distributions all provide very similar fits, with the GW marginally better than the lNpf, which in turn is marginally better than the GG.

We now consider an example using grouped data and no covariates.

6.2 *Bird lifetimes.*

Paranjpe and Rajarshi [11] fitted three parametric models to data from Deevey [4] on the age at death (grouped into one-year classes) for five bird species. They compared the fit of the Weibull distribution with what they called the *exponential* and the *double exponential power* models (these latter two both being capable of having bathtub-shaped hazards), using the Pearson χ^2 statistic, and showed that both of these two-parameter models fitted better than the Weibull, and yielded bathtub-shaped fitted hazards.

The lNpf was fitted to these five datasets along with another dataset from Deevey (lapwings), by maximizing the log-likelihood

$$\ell(\mu, \sigma, \beta) = \sum f_i \log[S_Y(\log i) - S_Y(\log i + 1)] \qquad (14)$$

where f_i is the number dying at age i. The estimates and χ^2 values are shown in Table 1 and the fitted hazard-rate functions are shown in Figure 5. Also shown are the χ^2 values for the generalized gamma (GG) fit. The dgrees of freedom column (df) applies to both the lNpf and GG fits. For all species, except starlings, the lNpf χ^2 values were smaller than those for the exponential power models (albeit on one fewer degree of freedom). For starlings the lNpf χ^2 value was between that for the exponential and double exponential power model. The AIC values (and χ^2 values)

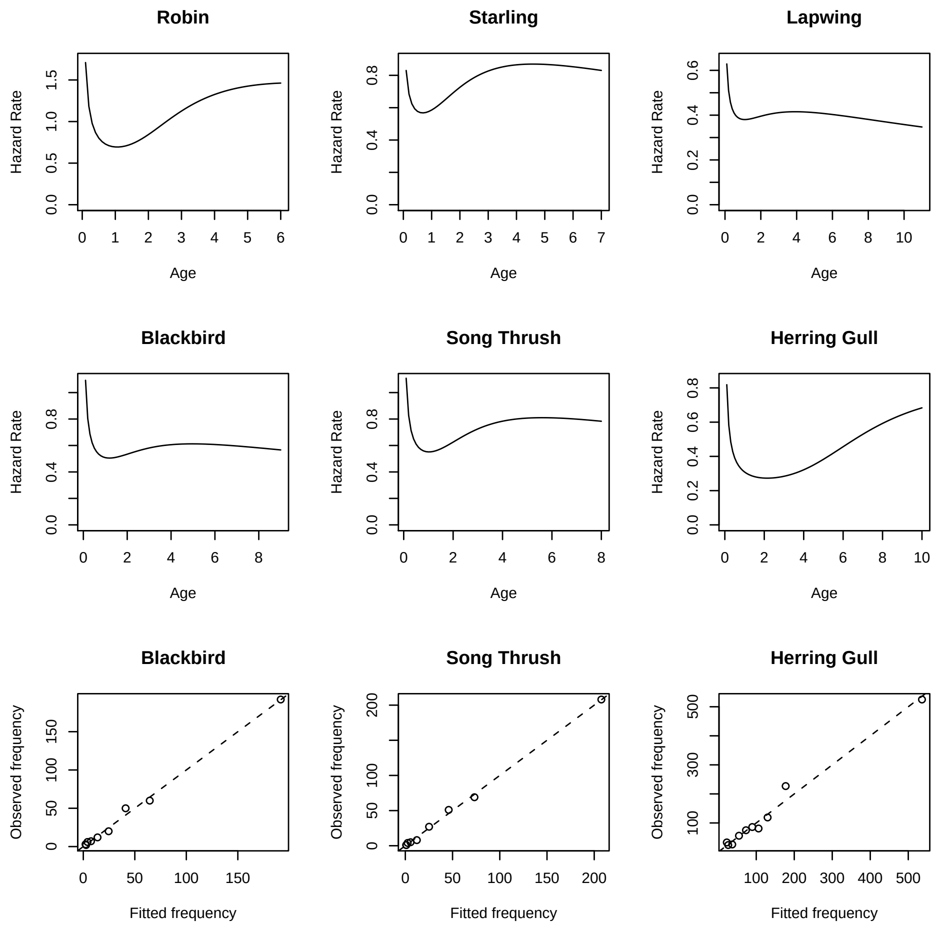

Figure 5. *Example 2. Bird lifetimes.* The top two rows show the fitted (MLE) hazard rates for six bird species. The bottom row shows observed *vs.* estimated frequencies for the three species in the second row. These include the best-fitting (song thrush) and worst fitting (herring gull).

Table 1. Maximum likelihood estimates (asymptotic standard errors in parentheses) of lNpf parameters for six bird species. Also shown (right-hand columns) are the χ^2 goodness-of-fit statistics for both the lNpf and GG models along with the corresponding (common) degrees of freedom.

Species	$\hat{\mu}$ (st. error)	$\hat{\sigma}$ (st. error)	$\hat{\beta}$ (st. error)	χ^2(lNpf)	(df)	χ^2(GG)
Robin	1.23 (.145)	0.295 (.106)	0.267 (.058)	3.07	(2)	3.62
Starling	1.22 (.152)	0.412 (.108)	0.613 (.123)	4.05	(3)	3.19
Lapwing	1.65 (.110)	0.555 (.080)	0.617 (.075)	6.76	(8)	5.68
Blackbird	1.47 (.110)	0.457 (.080)	0.426 (.054)	3.42	(5)	4.84
Song thrush	1.37 (.092)	0.398 (.067)	0.442 (.052)	3.41	(5)	5.01
Herring gull	2.07 (.039)	0.264 (.039)	0.412 (.019)	29.37	(6)	26.30

for the GG model (also with three parameters) were very close to those for the lNpf models (see Table 1), with the lNpf fitting better for robins, blackbirds and songthrushes, and the GG fitting better for the other three species.

The only species showing a significant lack of fit (of both lNpf and GG models) was the herring gull, for which the exponential power models had an even worse fit.

In the bottom row of Figure 5 plots of observed *vs.* fitted frequencies using the lNpf model, for three species (blackbird, songthrush and herring gull) are displayed. These include the best fitting (as determined by the observed χ^2 significance level) case (thrush) and the worse fitting case (herring gull).

In the next example the MLE of the lNpf distribution is a power function distribution (no lognormal component).

6.3 *Lifetime of devices on test*

Aarset [1] presented data on the lifetimes of 50 devices put on test and demonstrated that the hazard-rate function was bathtub shaped. There were no censored observations and no covariates. When fitting the lNpf to the data, the profile log-likelihood for σ was revealed to be increasing as $\sigma \to 0$. So for these data the likelihood is maximized on the boundary, $\sigma = 0$, and the best-fitting lNpf model is in fact simply a power-function distribution (*i.e.* a lNpf distribution with no log-normal component). When there is no censoring the MLEs of the two parameters μ and β of the power function distribution can be found analytically. As with the bathtub form of the generalized Weibull (GW) model [10] the MLE of the hazard rate is infinite at the largest failure time. The MLEs of the parameters of the power function distribution are:

$$\hat{\mu} = \log t_{(n)}; \qquad \hat{\beta} = \frac{n}{\sum_{i=1}^{n} \log(t_{(n)}/t_i)}$$

where $t_{(n)}$ is the largest failure time and, as with the GW model, the standard ML asymptotics do not hold. For the Aarset data $\hat{\mu} = 4.454$ and $\hat{\beta} = 0.727$ and the maximized log-likelihood is -219.89. In comparison the GW model, with three parameters, yielded a maximized log-likelihood of -218.07; while the generalized gamma (GG) a maximized log-likelihood of -220.03. The AIC values are 443.77 (power-function) and 442.13 (GW) and 446.06 (GG) so in terms of the AIC the GG model provides the worst fit and the GW a somewhat better fit than the power function model.

We now consider an example with covariates

6.4 *Steel under stress.*

McCool [9] gives the failure times for hardened steel specimens in a rolling contact fatigue test. Ten independent failure times were observed at each of four contact stresses. In an exercise Lawless [7, p.339] asks readers to assess a “Weibull power-law model” for these data *i.e.* a model in which the covariate stress (x) affects only the scale parameter of a Weibull distribution via a power-law relationship or in other words that the the hazard rate function is of Weibull form with scale parameter $\alpha = ax^b$ and unvarying shape-parameter. We consider fitting a model of this kind, using the lNpf distribution in place of the Weibull, *i.e.* assuming the covariate affects the hazard rate only by a change of scale. Since $\alpha = e^{\mu}$ is a scale-parameter for the lNpf model, this means fitting a model of the form

$$Y|x = \gamma_0 + \gamma_1 \log x + \sigma Z - \frac{1}{\beta}E \tag{15}$$

where $\gamma_0 = \log a$ and $\gamma_1 = b$. This is simply an accelerated life (AL) model with covariate $\log x$. Figure 6 shows a plot of failure time (log scale) against stress (log

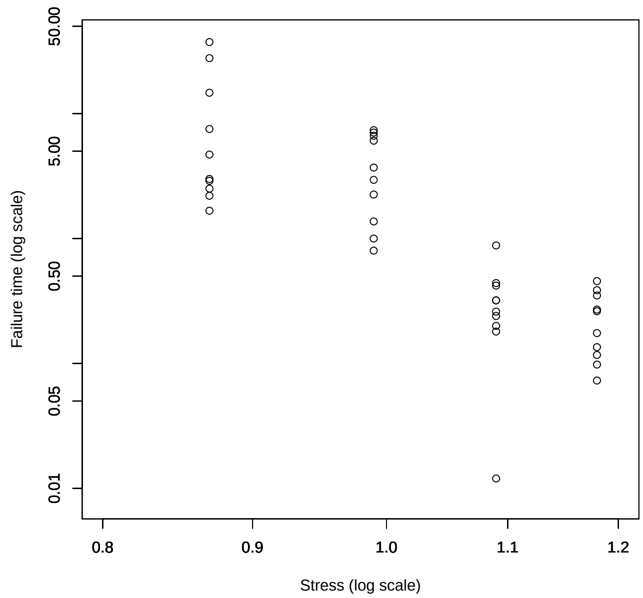

Figure 6. *Example 4. Steel under stress.* The failure time (log scale) *vs.* stress (log scale) for the data presented in Lawless (1982, p.339). Note the apparent outlier in the third group.

scale).

Fitting this model leads to ML estimates $(\hat{\gamma}_0, \hat{\gamma}_1, \hat{\sigma}, \hat{\beta}) = (13.51, -12.51, 0.772, 1.54)$ with asymptotic standard errors $(1.492, 1.370, 0.146, 0.579)$. The maximized log-likelihood is -56.62. If instead four separate lNpf models are fitted at each level of stress, the corresponding maximized log likelihood is -45.45. This leads to very strong evidence (P=.004) against the AL model (15). However there is an apparent outlier in the data (group 3). Refitting the AL model without this observation leads to a maximum of the likelihood on the boundary ($\beta = \infty$) of parameter space *i.e.* the best-fitting lNpf model is, in fact, one with a lognormal distribution with parameters $(\hat{\gamma}_0, \hat{\gamma}_1, \hat{\sigma}) = (0.342, -12.04, 0.749)$, with standard errors $(0.1412, 1.208, 0.170)$. There is no evidence (P=0.13) against the lognormal AL regression model, when the outlier is removed. Also without the outlier the lognormal regression model has a better fit (with AIC=105.40) than the Weibull regression model (AIC=114.32). Fitting the generalized gamma (GG) regression model, without the outlier, using the Prentice [[12]] parameterization again leads to a maximum of the likelihood on the boundary of parameter space corresponding once more to the lognormal model. Thus the performance of the lNpf and the GG are the same in this case.

When the outlier is included in the analysis the Weibull model (AIC=114.78) provides a better fit than the lNpf model (AIC=116.53), the generalized gamma model (AIC=121.16) and the lognormal model (AIC=122.09)

In the final example we consider the four-parameter extended lNpf model (11) with the proportional hazards assumption.

6.5 *Toxin-exposed and control groups of rats.*

Lagakos and Louis [6] present survival data for one hundred rats, half of which were insulted with 60 mg/kg of toluene diisocyanate, while the other half were kept as controls. Forty-four of the rats were still alive after 108 weeks. The extended lNpf

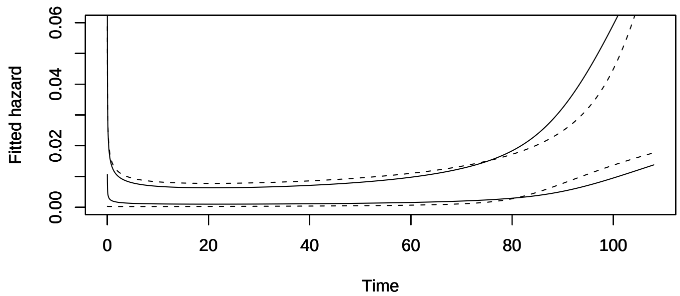

Figure 7. *Example 5. Toxin-exposed and control groups of rats.* The figure shows the fitted (MLE) hazard rates using the extended lNpf model (Section 4) for the two groups under the proportional hazards (PH) assumption (solid lines) and without this constraint (dashed lines).

model (11) was fitted to these data, first separately to each group and then to the combined group, assuming proportional hazards (*i.e* assuming common values to the parameters μ, σ and β, but possibly different values of θ for each group). Figure 7 shows the fitted hazards in the two cases, with the solid lines corresponding to proportional hazards, and dashed lines corresponding to separate fits. The likelihood ratio statistic for testing the PH assumption is not significant having a value 2.93 on 3 degrees of freedom. The MLE of the relative risk associated with the toxin is 6.24 and an approximate 95% confidence interval is $(3.38, 11.53)$. The extended lNpf model (AIC = 617.04) provides a much better fit than the ordinary lNpf model (AIC=1046.51). The generalized gamma model has a similar fit to the latter (AIC = 1046.31).

7. Conclusions.

The lognormal-power function distribution as a model for failure-time data is quite flexible in that its hazard-rate function can exhibit a variety of shapes, including a bathtub shape. Furthermore it lends itself naturally to accelerated-life (AL) regression modelling, and can be viewed as a fairly straightforward extension of the lognormal distribution. Parameter estimation by maximum likelihood requires numerical optimization but no difficulties have been experienced with multiple maxima or with obtaining convergence. In the paper the distribution has been fitted to a number of survival-time datasets taken from the literature. The fit of the lNpf distribution has been compared with the generalized gamma (GG) distribution with which it shares many properties. In the examples considered there is very little difference in fit (as measured by AIC) for the two models. To date no datasets with many covariates have been fitted but there is no reason to expect difficulties for the lNpf distributiuon apart from those experienced in other AL regression models (*e.g.* collinearity *etc.*). One can conclude that the lNpf distribution provides a viable model worthy of inclusion with the customary parametric models used in applied work.

Because the haxard-rate function of the lNpf distribution exhibits bathtub shaped behaviour only over a limited range and eventually decays like that of the lognormal distribution in the upper tail, as always in fitting complex models, care should be taken to not extrapolate beyond the range of the data.

Reed [14] considered an extension of the double Pareto-lognormal distribution, which (in the log-scale) was called the generalized normal-Laplace (GNL) distribution. A question for future research is whether a one-sided version of the this distribution (*i.e.* the distribution which results when the standard exponential random variable E in (1) is replaced by a gamma random variable with unit scale) would provide a useful extension of the lNpf model used in this paper. One major difficulty is that no closed-form for the GNL distribution is known, and hence no closed form for the corresponding log-likelihood. ML estimation will require more subtle methods (*e.g.* the E-M algorithm) than direct maximization.

Another subject for future research is whether the technique used for extending the lognormal distribution applied in this paper (multiplying a lognormal random variable by a power-law distributed random variable) can be applied to other two-parameter survival distributions, such as the Weibull, gamma, log-logistic *etc.* Some progress in this direction is reported in [15].

Acknowledgements

Research supported by NSERC grant OGP 7252 and largely carried out at the Department of Statistics, University of Waikato, whose support and hospitality are gratefully acknowledged. In particular I would like to thank Dr. Murray Jorgensen at Waikato for his contributions.

References

[1] M. V. Aarset *How to identify a bathtub hazard rate.* IEEE Trans. Reliability R-36 (1987), pp. 106-108.
[2] C. Cox, H. Chu, M. Schneider & A Muñoz *Parametric survival analysis and taxonomy of hazard functions for the generalized gamma distribution.* Statist. Med. 26 (2007) pp. 4352-4374
[3] C. Cox *The generalized* F *distribution: An umbrella for parametric survival analysis.* Statist. Med. 27 (2008) pp. 4301-4312.
[4] E. S. Deevey Jr. *Life-tables for natural populations of animals.* Quart. J. Biology 22 (1947), pp. 283-314.
[5] *Bathtub and related failure rate characterizations.* J. Amer. Stat. Assoc. 371 (1980), pp. 667-672.
[6] S. W. Lagakos & T. A. Louis *Use of tumour lethality to interpret tumorigenicity experiments lacking cause-of-death data.* Appl. Statist. 37 (1988), pp. 169-179.
[7] J. F. Lawless *Statistical Models and Methods for Lifetime Data,* New York: John Wiley and Sons, 1982.
[8] A. W. Marshall & I. Olkin. *Life Distributions* New York: Springer, 2007.
[9] J. I. McCool *Confidence limits for Weibull regression with censored data.* IEEE Trans. Reliability R-29, (1980) pp. 145-150.
[10] G. S. Muldholkar, D. K. Srivastava & G. D. Kollia *A Generalization of the Weibull distribution with application to the analysis of survival data.* J. Amer. Stat. Assoc. 91, (1996) pp. 1575-1583.
[11] S. Paranjpe & M. B. Rajarashi *Modelling non-monotonic survivorship data with bathtub distributions.* Ecology 67, (1986) pp. 1693-1695.
[12] R. L. Prentice *A log gamma model and its maximum likelihood estimation.* Biometrika 61 (1974) pp 539-544
[13] W. J. Reed & M. Jorgensen *The double Pareto-lognormal distribution - A new parametric model for size distributions.* Comm. Stats - Theory & Methods, 33, (2004) pp. 1733-1753.
[14] W. J. Reed *Brownian-Laplace motion and its use in financial modelling.* Comm. Stats - Theory & Methods, 36 (2007) pp. 473-484.
[15] W. J. Reed *Power-law adjusted failure-time models.* Submitted to Comm. Stats - Theory & Methods, (2010).
[16] H. Ruben *A new asymptotic expansion for the normal probability integral and Mill's ratio.* J. Roy. Statist. Soc. B, 24, (1962), pp. 177-179.

Appendix A. Tail behaviour of hazard-rate function.

Upper tail. Using (7) and letting $z = (\log t - \mu)/\sigma$, one can write

$$h(t) = \beta e^{-\mu-\sigma z} \frac{R(\beta\sigma + z)}{R(z) - R(\beta\sigma + z)}. \tag{A1}$$

A well-known property of Mills' ratio is that $R(x) \sim x^{-1}$ as $x \to \infty$ (see *e.g.* Ruben, 1962). Using this in the above yields

$$h(t) \sim \beta e^{-\mu-\sigma z} \frac{z}{\beta\sigma} = \frac{\log t - \mu}{\sigma^2 t}$$

as $t \to \infty$, which is independent of β and is exactly the tail behaviour of the lognormal hazard with parameters μ and σ^2.

Lower tail. Write R_β for $R(\beta\sigma + (\log t - \mu)/\sigma)$ and R_0 for $R((\log t - \mu)/\sigma)$, so that the hazard rate (A1) can be expressed as

$$h(t) = \frac{\beta}{t} \frac{R_\beta/R_0}{[1 - R_\beta/R_0].} \tag{A2}$$

Now

$$\frac{R_\beta}{R_0} = \frac{\phi(z)}{\phi(z+\beta\sigma)} \left[\frac{\Phi^c(z+\beta\sigma)}{\Phi^c(z)}\right] = \exp(\beta\sigma z + \beta^2\sigma^2/2) \left[\frac{\Phi^c(z+\beta\sigma)}{\Phi^c(z)}\right] \tag{A3}$$

which tends to zero as $t \to 0$ ($z \to -\infty$) since the term in square brackets $\to 1$. So $h(t)$ behaves at zero like

$$\frac{\beta}{t}\frac{R_\beta}{R_0} = \beta \exp((\beta - 1)\sigma z - \mu + \beta^2\sigma^2/2) \left[\frac{\Phi^c(z+\beta\sigma)}{\Phi^c(z)}\right].$$

Since $z \to -\infty$ as $t \to 0$, it is clear that β is a critical parameter: if $\beta > 1$, the limiting value of h at $t = 0$ is 0; while if $\beta < 1$, it is ∞ (a vertical asymptote of h at 0); in the intermediate case ($\beta = 1$), $h(t) \to e^{-\mu+\sigma^2/2}$. Qualitatively this behaviour is similar to that of the power function hazard. However quantitatively it is different, being dependent on the parameter σ^2 of the lognormal component.

Appendix B. Shape of hazard-rate function.

One can write the representation (A2) of the hazard-rate function as

$$h(t) = \frac{\beta}{t} G(r(z(t))), \tag{B1}$$

say, where

$$z(t) = \frac{\log(t) - \mu}{\sigma}, \qquad r(z) = \frac{R_\beta(z)}{R_0(z)} \quad \text{and} \quad G(r) = \frac{r}{1-r} \tag{B2}$$

from which it follows by the chain rule that

$$h'(t) = \frac{\beta}{t^2} \frac{r(z)}{(1-r(z))^2} \left[\frac{1}{\sigma}\frac{r'(z)}{r(z)} - (1 - r(z))\right] \tag{B3}$$

so that

$$\mathrm{sgn}(h'(t)) = \mathrm{sgn}\left(\frac{1}{\sigma}\frac{r'(z)}{r(z)} - (1 - r(z))\right) = \mathrm{sgn}[F_1(z) - F_2(z)], \quad \text{say},$$

where

$$F_1(z) = \frac{1}{\sigma}\frac{r'(z)}{r(z)} = \beta - \frac{1}{\sigma}\left[\frac{1}{R_\beta} - \frac{1}{R_0}\right]$$

is a decreasing function, with left and right limits (as $z \to -\infty$ and $z \to \infty$, respectively) of β and 0; and

$$F_2(z) = 1 - r(z)$$

is a decreasing function with left and right limits 1 and 0.

The asymptotic behaviour of $F_1(z)$ and $F_2(z)$ as $z \to \infty$ can be determined from the Laplace asymptotic expansion of the Mills' ratio (see *e.g.* Ruben, 1962)

$$R(x) = \frac{1}{x}\left(1 - \frac{1}{x^2} + \frac{1.3}{x^4} - \frac{1.3.5}{x^6} + \ldots\right) \tag{B4}$$

One obtains

$$F_1(z) \sim \beta/z^2 \qquad \text{and} \qquad F_2(z) \sim \beta\sigma/z \qquad \text{as} \quad z \to \infty$$

From this it follows that $F_2(z) > F_1(z)$ for z suitably large.

If $\beta > 1$, it follows from the above that F_1 and F_2 cross for $y \in (\infty, \infty)$ an odd number of times, so that $h(t)$ has an odd number of turning points. In contrast if $\beta < 1$, F_1 and F_2 cross an even number of times, with $h(t)$ having an even number of turning points. In fact numerical explorations indicate that the number of crossings, for $\beta > 1$ is always one; and the number for $\beta < 1$ is zero or two.

Coupled with the earlier results on the tail behaviour of $h(t)$ this suggests that when $\beta > 1$ the hazard-rate $h(t)$ is unimodal, increasing at first from zero and subsequently decreasing asymptotically to zero; and when $\beta < 1$ the hazard-rate $h(t)$ either monotone decreasing (from infinity, at zero, to zero at infinity), or is at first decreasing (from $h(0) = \infty$), then increases to a local maximum, and subsequently decreases to 0 as $t \to \infty$. It is in this latter case that a bathtub shaped hazard can occur (over a limited range of t). Thus a necessary condition for a bathtub shaped hazard is that $\beta < 1$.

In the intermediate case ($\beta = 1$), the hazard-rate has zero slope at $t = 0$ and assumes the finite non-zero value value $e^{-\mu+\sigma^2/2}$.

For further analysis we note that F_1 can be written $F_1(z) = F_0(z)/\sigma$ where

$$F_0(z) = \left[\beta\sigma - \frac{1}{R_\beta} + \frac{1}{R_0}\right].$$

We now consider varying σ so that $\beta\sigma$ remains fixed at a value B, say (*i.e.* moving along the hyperbolic locus $\beta\sigma = B$, for a given B). For σ suitably large $F_1(z) = F_0(z)/\sigma$ is smaller than $F_2(z)$ everywhere on $(-\infty, \infty)$. In this case the hazard rate has no turning points and is monotone decreasing. As σ decreases one of two things can happen: either $\beta = B/\sigma$ can become greater than 1 ; or F_1 can increase and 'push through' F_2 so that it is larger than F_2 on a finite interval, with β still less than 1. In the former case there is a bifurcation from no turning points for the hazard to one turning point (hazard moves from monotone to unimodal); in the latter case there is a bifurcation from no turning points for the hazard to two turning points (hazard moves from monotone to having one local minimum followed by one local maximum). It follows that σ-β parameter space can be divided into three disjoint regions corresponding to 0, 1 and 2 turning points for the hazard rate (see Figure 2).

To determine the locus of points where the bifurcation from 0 to 2 turning points occurs, one can consider for fixed B the solution (in z and σ) to the pair of equations

$$\sigma^{-1}F_0(z) = F_1(z) \quad \text{and} \quad \sigma^{-1}F_0'(z) = F_1'(z)$$

(where the prime denotes a derivative), because for these values of σ and z, F_1 is tangent to F_2. Dividing the above two equations leads to the single equation for z

$$\frac{F_0'(z)}{F_0(z)} = \frac{F_2'(z)}{F_2(z)}$$

which can be expressed in terms of the Mills ratio, R, and its derivative. Using the fact that $R'(x) = xR(x) - 1$ simplifies the equation to one in z involving only R. This equation can be solved numerically. From this solution ($z^\star$, say) the value $\sigma^\star$ (and $\beta^\star = B/\sigma^\star$) at which the bifurcation occurs can be obtained (as $F_0(z^\star)/F_2(z^\star)$). Repeating this for various values of B yields the locus of points $(\sigma^\star, \beta^\star)$ on the bifurcation boundary. The curve separating Regions 2 and 0 in Figure 2 was calculated in this way.

Weibull-like behaviour of lNpf model.

In this section we show analytically that the lNpf distribution is approximated by the Weibull distribution for $t << e^\mu$, when $\beta\sigma$ is small.

We consider the hazard-rate for $y = \log t$, which from (7) can be written

$$h_Y(y) = \frac{\beta}{(R_0/R_\beta - 1)}$$

Now using (A3) and expanding $\Phi^c(z+\beta\sigma)$, assuming $\beta\sigma$ is small, the ratio R_0/R_β can be written

$$\frac{R_0}{R_\beta} = \exp(-\beta\sigma z)\frac{1}{1-\beta\sigma/R(z)} + o(\beta\sigma)$$

and so

$$h_Y(y) = \frac{\beta(1-\beta\sigma/R(z))}{e^{-\beta\sigma z} - 1 + \beta\sigma/R(z)} + o(\beta\sigma).$$

Now for t sufficiently small (so that $\log t << \mu$), z will be large negative and $R(z)$ large positive. Coupled with the fact that we assume $\beta\sigma$ small, it follows that for t suitably small $\beta\sigma/R(z)$ will be very small and so approximately

$$h_Y(y) \approx \frac{\beta e^{\beta y}}{e^{\beta\mu} - e^{\beta y}}$$

which is the form of the hazard (in $Y = \log T$) for the power-function distribution (see (9)), with

$$Y = \log T = \mu - \frac{1}{\beta}E,$$

so the power-function component of the distribution predominates in determining the hazard at the lower end. The cumulative hazard $H_Y(y) = -\log S_Y(y)$ for this power-function distribution equals $-\log(1 - e^{\beta(y-\mu)})$, for $y < e^\mu$ and is infinite otherwise. So for $y << e^\mu$, $H_Y(y) \approx e^{\beta(y-\mu)}$, and $\log H_Y(y) \approx \beta(y - \mu)$, which is linear in y. This is exactly the cumulative hazard for $\log T$ when T has the Weibull survivor function $\exp(-e^{-\beta\mu} t^\beta)$. Thus when μ is large and $\beta\sigma$ small, the lNpf distribution should be well-approximated by the Weibull distribution, at least for $t << e^\mu$.