Survival distributions based on the XD model class

Kevin Burke[1] and Gilbert MacKenzie[1,2]

[1] Centre for Biostatistics, University of Limerick, Ireland.
[2] ENSAI, Rennes, France.

Email: kevin.burke@ul.ie gilbert.mackenzie@ul.ie

Abstract: We aim to explore the survival distributions based on the extreme dispersion, XD, models proposed by Jørgensen (2010). It is suggested that survival times can be modelled within the XD framework by taking $\log T = Y \sim XD(\mu, \lambda)$, where T is the survival time and Y is the extreme dispersion random variable. We will show how these survival models can be generated using XD and look at the distribution and properties of the survival time $T = e^Y$.

Keywords: Extreme dispersion; Survival; Morris class; Quadratic slope.

1 Introduction

Jørgensen (2010) proposed an extreme value analogues of exponential dispersion models and generalized linear regression models. The slope function is introduced as an analogy to the variance function. Therefore the slope function characterizes the extreme dispersion model in much the same way that the variance function characterizes the exponential dispersion model. To begin the analogy, we know that the moment generating function is $M(t) = E(e^{tY})$ and the cumulant generating function is $K = \log M$. The first two derivatives of the c.g.f. evaluated at $t = 0$ are the mean and variance. Defining $\tau = K'$, *the mean value mapping*, we have,

$$E(Y) = \tau(0), \qquad Var(Y) = \tau'(0). \tag{1}$$

It should be noted that τ is strictly increasing.
In the XD framework the survivor function $\Pr(Y > y) = G(y)$, where $Y \in \mathcal{C}$, is analogous to the moment generating function. Therefore, apart from a sign change, the cumulative hazard function $H = -\log G$ is analogous to the cumulant generating function. As we know, $h = H'$ is the hazard function and so we have

$$r(Y) = h(0), \qquad s(Y) = h'(0), \tag{2}$$

where r is the *rate* and s is the *slope*. Unlike the variance, the slope can be negative as well as positive. The hazard function, h, must be monotone like the mean value mapping is monotone. However, h can be monotone increasing or decreasing. Interestingly, Jørgensen (2010) shows that the hazard function for $T = e^Y$ need not have monotone hazard rate, even if Y does. Therefore we are not limited in this sense when considering the survival extension of the XD model.
Going back to the random variable, Y, the *rate* is denoted $\mu \in \Psi = h(\mathcal{C})$ and the *unit slope function* is defined as

$$v(\mu) = h'(h^{-1}(\mu)) \tag{3}$$

which maps Ψ onto $\mathcal{R}_+$ when h is increasing and onto $\mathcal{R}_-$ when h is decreasing. Here $h^{-1}(\cdot)$ is the inverse hazard function so that $h(h^{-1}(\mu)) = \mu$.
It can be shown that the inverse hazard function satisfies the differential equation

$$\frac{\mathrm{d}h^{-1}(\mu)}{\mathrm{d}\mu} = \frac{1}{v(\mu)}. \tag{4}$$

Replacing h with τ and v with V in equations 3 and 4 gives us the equations in the exponential dispersion framework which show the relationship between the mean value mapping and the unit variance function.

2 The hazard location family

The one parameter hazard location family $HL(\mu)$ is introduced as a precursor to the XD model is given by the family of models with survivor function

$$G(y; \mu) = G(y + h^{-1}(\mu)) \tag{5}$$

with support $\mathcal{C} - h^{-1}(\mu)$ and $\mu \in \Psi$.
So to generate the HL family we start with a survivor function $G(y)$ and work out the corresponding inverse hazard function which is parametrized by μ. We then replace y by $y + h^{-1}(\mu)$ in $G(y)$. This gives us the survivor function for the HL family, $G(y; \mu)$. Note that the survivor function $G(y)$ has no parameters; $G(y)$ is used merely to generate $G(y; \mu)$ which has one parameter.
The hazard location family is analogous to the natural exponential families. Natural exponential families are the one parameter exponential models that either have no dispersion parameter or the dispersion parameter is known.

3 The extreme dispersion model

The extreme dispersion model, $XD(\mu, \lambda)$, generated by G has survivor function

$$G(y; \mu, \lambda) = G^{\lambda}\left(\frac{y}{\lambda} + h^{-1}(\mu)\right) \tag{6}$$

which has support on $\mathcal{C}^* = \lambda(\mathcal{C} - h^{-1}(\mu))$. Here $\mu \in \Psi$ and $\lambda > 0$. So to generate the XD survivor function we replace y by $\frac{y}{\lambda} + h^{-1}(\mu)$ in $G(y)$ and raise it to the power of λ.

The reason for carrying out this operation is based on properties of the moment generating function and the fact that $G(y)$ is analogous to this function.

The cumulative hazard function for the XD model is

$$H(y; \mu, \lambda) = \lambda H\left(\frac{y}{\lambda} + h^{-1}(\mu)\right), \tag{7}$$

where $H(\cdot) = -\ln G(\cdot)$. And the hazard function is

$$h(y; \mu, \lambda) = h\left(\frac{y}{\lambda} + h^{-1}(\mu)\right) \tag{8}$$

where $h(\cdot) = H'(\cdot)$.

The density function for the XD model is therefore

$$f(y; \mu, \lambda) = h\left(\frac{y}{\lambda} + h^{-1}(\mu)\right) \exp\left(-\lambda H\left(\frac{y}{\lambda} + h^{-1}(\mu)\right)\right). \tag{9}$$

It can be shown that $h(0; \mu, \lambda) = \mu$ is the rate and $h'(0; \mu, \lambda) = \frac{1}{\lambda} v(\mu)$ is the slope for the XD model.

4 Generators: the Morris class

Jørgensen (2010) puts special emphasis on XD models with quadratic slope function. These are analogous to the exponential dispersion models with quadratic variance functions known as the Morris class. The table below gives the relevant functions needed to generate the $XD(\mu, \lambda)$ model for the seven cases listed in Jørgensen (2010).

5 XD survival model

To obtain the survival model based on the XD model, we let $Y = \log T$ where T is the positive survival time. This gives us the following

$$G_T(t; \mu, \lambda) = G^{\lambda}\left(\frac{\ln t}{\lambda} + h^{-1}(\mu)\right), \tag{10}$$

TABLE 1. *Generators*

	$G(y)$	$H(y)$	$h(y)$	$h^{-1}(\mu)$	$\mathcal{C}$	$v(\mu)$	$\psi = h(\mathcal{C})$
Rayleigh	$e^{-\frac{y^2}{2}}$	$\frac{y^2}{2}$	y	μ	$\mathcal{R}_+$	1	$\mathcal{R}_+$
Gumbel	e^{-e^y}	e^y	e^y	$\ln \mu$	$\mathcal{R}$	μ	$\mathcal{R}_+$
Uniform	$1-y$	$-\ln(1-y)$	$\frac{1}{1-y}$	$1-\frac{1}{\mu}$	$(0,1)$	μ^2	$(1,\infty)$
Pareto	$\frac{1}{y}$	$\ln y$	$\frac{1}{y}$	$\frac{1}{\mu}$	$(1,\infty)$	$-\mu^2$	$(0,1)$
Logistic	$\frac{1}{1+e^y}$	$\ln(1+e^y)$	$\frac{e^y}{1+e^y}$	$\ln \frac{\mu}{1-\mu}$	$\mathcal{R}$	$\mu(1-\mu)$	$(0,1)$
Neg Exp	$1-e^y$	$-\ln(1-e^y)$	$\frac{e^y}{1-e^y}$	$\ln \frac{\mu}{1+\mu}$	$\mathcal{R}_-$	$\mu(1+\mu)$	$\mathcal{R}_+$
Cosine	$\cos y$	$-\ln(\cos y)$	$\tan y$	$\tan^{-1} \mu$	$(0,\frac{\pi}{2})$	$1+\mu^2$	$\mathcal{R}_+$

$$H_T(t;\mu,\lambda) = \lambda H\left(\frac{\ln t}{\lambda} + h^{-1}(\mu)\right), \tag{11}$$

$$h_T(t;\mu,\lambda) = \frac{1}{t} h\left(\frac{\ln t}{\lambda} + h^{-1}(\mu)\right), \tag{12}$$

and,

$$f_T(t;\mu,\lambda) = \frac{1}{t} h\left(\frac{\ln t}{\lambda} + h^{-1}(\mu)\right) \exp\left(-\lambda H\left(\frac{\ln t}{\lambda} + h^{-1}(\mu)\right)\right), \tag{13}$$

where the subscript T here indicates that these functions correspond to the survival time T.

6 Estimation

Jørgensen (2010) suggests a quasi-likelihood method for fitting these models (in the regression setting). He also mentions maximum likelihood as a possibility and discusses some potential problems with both methods. Here we will look at the latter estimation method, maximum likelihood. One possible problem with this method is the fact that for many of the distributions (those where $\mathcal{C} \neq (-\infty,\infty)$) the parameters μ and λ will appear in the support. This can cause problems with maximum likelihood and, more specifically, the proof of unbiasedness of of MLEs in maximum likelihood theory requires that the parameters being estimated do not appear in the support.

However this needs more investigation and for now we will proceed with this method. Looking only at the survival XD model, the likelihood function is given by:

$$L_T(\mu,\lambda) = \prod_{i=1}^{n} \frac{1}{t_i} h\left(\frac{\ln t_i}{\lambda} + h^{-1}(\mu)\right) \exp\left(-\lambda H\left(\frac{\ln t_i}{\lambda} + h^{-1}(\mu)\right)\right), \tag{14}$$

which can be extended to include censoring in the usual way,

$$L_T(\mu,\lambda) = \prod_{i=1}^{n} \left[\frac{1}{t_i} h\left(\frac{\ln t_i}{\lambda} + h^{-1}(\mu)\right)\right]^{\delta_i} \exp\left(-\lambda H\left(\frac{\ln t_i}{\lambda} + h^{-1}(\mu)\right)\right), \tag{15}$$

where δ_i is the censoring indicator.
Thus, the log-likelihood is

$$\ell_T(\mu,\lambda) = \sum_{i=1}^{n} \delta_i \log\left[\frac{1}{t_i} h\left(\frac{\ln t_i}{\lambda} + h^{-1}(\mu)\right)\right] - \lambda H\left(\frac{\ln t_i}{\lambda} + h^{-1}(\mu)\right). \tag{16}$$

7 Worked Example: Rayleigh-XD

We now look at a worked example; the "Rayleigh-XD", so called because it is the $XD(\mu,\lambda)$ model based on the Rayleigh generator. Using the relevant functions in the first row in Table 1 and the equations in Sections 3 and 5, we can obtain the XD model and it's survival counterpart.

7.1 XD Model

$$G(y;\mu,\lambda) = \exp\left(-\frac{\lambda}{2}\left(\frac{y}{\lambda} + \mu\right)^2\right) \tag{17}$$

$$H(y;\mu,\lambda) = \frac{\lambda}{2}\left(\frac{y}{\lambda} + \mu\right)^2 \tag{18}$$

$$h(y;\mu,\lambda) = \frac{y}{\lambda} + \mu \tag{19}$$

$$Support : \mathcal{C}^* = \lambda \times ([0,\infty) - \mu) = [-\lambda\mu, \infty) \tag{20}$$

7.2 Survival

$$G_T(t;\mu,\lambda) = \exp\left(-\frac{\lambda}{2}\left(\frac{\log t}{\lambda} + \mu\right)^2\right) \tag{21}$$

$$H_T(t;\mu,\lambda) = \frac{\lambda}{2}\left(\frac{\log t}{\lambda} + \mu\right)^2 \tag{22}$$

$$h_T(t;\mu,\lambda) = \frac{1}{t}\left(\frac{\log t}{\lambda} + \mu\right) \tag{23}$$

$$Support : \mathcal{C}_T = \exp(\mathcal{C}^*) = [\exp(-\lambda\mu), \infty) \tag{24}$$

8 Simulation

We simulated survival times from the survival Rayleigh-XD. We did this for all combinations of $\mu = (2, 0.5)$, $\lambda = (2, 0.5)$ and $n = (100, 1000)$, where n is the sample size. So in total there were $2 \times 2 \times 2 = 8$ different settings. There was no censoring for the purposes of this simulation. Survival times were generated using the inverse function,

$$t = F^{-1}(u) = \exp\left(\lambda\left\{\left[-\frac{2}{\lambda}\log(1-u)\right]^{1/2} - \mu\right\}\right). \tag{25}$$

Maximum likelihood was then used to fit the survival Rayleigh-XD model to the simulated data. Within each setting this was done 1000 times. The average of the 1000 MLEs for each setting is shown in the Table 2 below, along with the % bias given by $pbias = 100(\tilde{\theta} - \theta)/\theta$, where $\tilde{\theta}$ represents the average of the 1000 MLEs.

TABLE 2. Simulation Results

	μ	λ	n	$\tilde{\mu}$	$pbias_\mu$	$\tilde{\lambda}$	$pbias_\lambda$
1	2.0	2.0	100	2.07	3.54	1.94	-2.80
2	0.5	2.0	100	0.50	0.58	1.95	-2.53
3	2.0	0.5	100	2.05	2.69	0.49	-2.42
4	0.5	0.5	100	0.49	-2.86	0.49	-2.40
5	2.0	2.0	1000	2.01	0.51	1.99	-0.53
6	0.5	2.0	1000	0.50	-0.02	1.99	-0.54
7	2.0	0.5	1000	2.01	0.34	0.50	-0.46
8	0.5	0.5	1000	0.50	-0.51	0.50	-0.44

We can see from the Table 2 that there does not seem to be an issue with the MLEs in this simple setting for the survival Rayleigh-XD.

9 Discussion

The XD class of extreme dispersion models offers the prospect of developing a new class of survival models with novel properties. Our early work confirms that this is indeed the case. However, the resulting survival models appear to embody a latent period before which failure becomes operational. Whilst this is always a testable hypothesis, in the group setting, the extent to which these models can be usefully applied in practice remains to be ascertained.

Acknowledgments

This work was supported, in part, by the SFI's (www.sfi.ie) BIO-SI research programme, grant number, **07MI012**. The first author is an IRCSET Scholar (www.ircset.ie) and the second is the Principal Investigator of BIO-SI (www.ul.ie/bio-si).

References

Coles, S.G. (2001) *An Introduction to Statistical Modeling of Extreme Values.* Springer.

Collett, D. (2003) *Modelling Survival Data in Medical Research, 2nd ed.* Chapman & Hall/CRC.

Jorgensen, B. (1986) Some properties of exponential dispersion models. *Scand. J. Statist.*, **13**, 187-197.

Jorgensen, B. (1987) Exponential dispersion models (with discussion). *J. R. Stat. Soc.*, **49**, 127-162.

Jorgensen, B. (1997) *The Theory of Dispersion Models.* Chapman & Hall.

Jorgensen, B., Goegebeur, Y. and Martinez, J.R. (2010) Dispersion models for extremes. *Extremes*, **13**, 399-437.

McCullagh, P., Nelder, J.A. (1989) *Generalized Linear Models, 2nd ed.* Chapman & Hall.