LIFETIME MODELS WITH NONCONSTANT SHAPE PARAMETERS

Authors: Josmar Mazucheli
– Departamento de Estatística, Universidade Estadual de Maringá, Maringá, P.R. — Brazil (jmazucheli@uem.br)

Francisco Louzada-Neto
– Departamento de Estatística, Universidade Federal de São Carlos, São Carlos, S.P. — Brazil (dfln@power.ufscar.br)

Jorge Alberto Achcar
– Departamento de Estatística, Universidade Federal de São Carlos, São Carlos, S.P. — Brazil (jorge@icmc.sc.usp.br)

Received: October 2002 Revised: September 2003 Accepted: October 2003

Abstract:

- In its standard form, a lifetime regression model usually assumes that the time until an event occurs has a constant shape parameter and a scale parameter that is a function of covariates. In this paper we consider lifetime models with shape parameter dependent on a vector of covariates. Two special models are considered, the Weibull model and a mixture model incorporating long-term survivors, when we consider that the incidence probability is also dependent on covariates. Classical parameters estimation approach is considered on two real data sets.

Key-Words:

- *accelerated life tests; bootstrap; long-term survivors; nonconstant shape parameter; Weibull distribution.*

AMS Subject Classification:

- 49A05, 78B26.

1. INTRODUCTION

To express the distribution of a nonnegative random variable, T, which represents the lifetime of individuals (or components) in some population subjected to covariate effects, several mathematically equivalent functions that uniquely determine the distribution can be considered; namely, the cumulative distribution, the density, the survival and the hazard functions [16]. For lifetime data, the survival function at a particular time t is defined as

$$S_0(t \,|\, \mu(\mathbf{x}), \gamma) \;=\; P(T > t \,|\, \mu(\mathbf{x}), \gamma) \;, \tag{1.1}$$

where $\mu(\mathbf{x})$ is a scale parameter that is a function of covariate involving unknown parameters and γ is a constant unknown shape parameter. It is particularly useful to define the survival model in terms of (1.1), because of its interpretation as the probability of an individual (or component) surviving till time t [16].

Besides, in several applications, it is clear that a non-zero proportion of patients or components can be considered cured, or do not fail in their testing time [20]. In this context, we consider the model

$$S(t \,|\, \mathbf{x}) \;=\; p + (1-p)\, S_0(t \,|\, \mu(\mathbf{x}), \gamma) \;, \tag{1.2}$$

where S is the population survival function and $0 < p < 1$ represents the cured fraction, which is cured or never fails with respect to the specific cause of death (or failure). Observe that (1.2) is a mixture model with two components, where S_0 is the survival function of the individuals which are not cured. For the cured patients, the survival function is equal to one for all finite values t. Mixture survival models provide a way of modelling time to death when cure is possible, simultaneously estimating death hazard of fatal cases and the proportion of cured cases.

In many applications however the usual assumption of constant shape parameter γ cannot be appropriate. For instance, in some studies with fatigue of materials, usually, it is assumed that the shape parameter of the Weibull distribution depends on the stress levels, as we can see in Wang and Kececioglu ([30]), Meeker and Escobar ([22]), Pascual and Meeker ([26]), Meeter and Meeker ([23]), Meeker and Escobar ([21]), Hirose ([12]), Chan ([4]), Smith ([27]) and Nelson ([25]). Anderson ([1]) considers a Weibull accelerated regression model with the dispersion parameter depending on the location parameter. In the context of risk modelling, Hsieh ([13]) introduces heteroscedastic risk models, and Louzada-Neto ([19, 17]) introduces an extended risk model. Applications in the context of regression models with normal errors and nonconstant scale are considered by Zhou *et al.* ([31]) and Tanizaki and Zhang ([28]). Cepeda and Gamerman ([3]) consider Bayesian modelling of variance heterogeneity in normal regression models.

In this paper we consider a general survival model with shape and cured fraction parameters depending on covariates. The approach with constant shape parameter was first used by Farewell [8]. The advantage of such a formulation

is to have several usual survival models as particular cases. Maximum likelihood estimation procedure is adopted for two special cases: the Weibull distribution with shape parameter depending on a vector of covariates and a long-term Weibull survival mixture model in the presence of covariates. In Section 2 we introduce a general survival model with shape and scale parameters depending on covariates. The Weibull case is introduced in Section 3. Two real data sets are presented in Section 4. Some concluding remarks in Section 5 finalize the paper.

2. A GENERAL SURVIVAL MODEL

Consider a survival model with shape parameter depending on covariates. The corresponding survival function is

$$S_0(t \,|\, \mu(\mathbf{x}), \gamma(\mathbf{y})) \,=\, P(T > t \,|\, \mathbf{x},\, \mathbf{y}) \;, \tag{2.1}$$

where $\mu(\mathbf{x})$ is a scale parameter depending on a covariate vector, $\mathbf{x}$, and $\gamma(\mathbf{y})$ is the shape parameter depending on a covariate vector, $\mathbf{y}$. Both $\boldsymbol{\mu}$ and $\boldsymbol{\gamma}$ may involve unknown parameters. Of course, the vectors $\mathbf{x}$ and $\mathbf{y}$ can be equal.

For fitting long-term survival data, where a proportion of the individuals are cured [20], we consider the general survival model

$$S(t|\mathbf{x}, \mathbf{y}, \mathbf{z}) \,=\, p(\mathbf{z}) + (1-p(\mathbf{z}))\, S_0(t \,|\, \mu(\mathbf{x}), \gamma(\mathbf{y})) \;, \tag{2.2}$$

where $\mu(\mathbf{x})$ and $\gamma(\mathbf{y})$ are scale and shape parameters of the lifetime distribution of non-cured patients and $0 < p(\mathbf{z}) < 1$ is the incidence probability depending on a covariate vector, $\mathbf{z}$, involving unknown parameters. For $p(\mathbf{z}) = 0$ we have the model (2.1).

A special case is given by the Weibull survival function for the non-cured patients, given by

$$S_0(t \,|\, \mu(\mathbf{x}), \gamma(\mathbf{y})) \,=\, \exp\left[-\left(\frac{t}{\mu(\mathbf{x})}\right)^{\gamma(\mathbf{y})}\right]. \tag{2.3}$$

Let us assume a random sample $T_1, ..., T_n$, such that, associated to each T_i there are covariate vectors $\mathbf{x}_i^t = (1, x_{i1}, ..., x_{ik})$, $\mathbf{y}^t = (1, y_{i1}, ..., y_{ik})$ and $\mathbf{z}^t = (1, z_{i1}, ..., z_{ik})$, and an indicator variable δ_i, $\delta_i = 1$ if t_i is an observed lifetime or $\delta_i = 0$ if t_i is a censored observation (rigth-censored observations). Then, for an uninformative censoring mechanism, the likelihood function [16] can be written as

$$L \,=\, \prod_{i=1}^{n} f(t_i|\mathbf{x}_i, \mathbf{y}_i, \mathbf{z}_i)^{\delta_i}\, S(t_i|\mathbf{x}_i, \mathbf{y}_i, \mathbf{z}_i)^{1-\delta_i} \;, \tag{2.4}$$

where $f(t_i|\mathbf{x}_i, \mathbf{y}_i, \mathbf{z}_i)$ is the density function and $S(t_i|\mathbf{x}_i, \mathbf{y}_i, \mathbf{z}_i)$ is defined in (2.2).

Let $\boldsymbol{\theta}^{'} = (\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma})$ be the parameter vector indexing (2.2). The maximum likelihood estimator (MLE) of $\boldsymbol{\theta}$ can be obtained by solving the system of nonlinear equations, $\partial \log L/\partial \boldsymbol{\theta} = \mathbf{0}$. However, it can be hard to solve the system of nonlinear equations above by pure Newton-type methods, since it is easy to overstep the true minimum. An alternative algorithm is proposed by [30] based on [15, 2, 9]. However, a straightforward procedure, which we prefer, is to maximize (2.4). This procedure can be implemented in a standard statistical software such as *R* [14] or a SAS via a routine that finds a local maximum of a nonlinear function using general-purpose optimization procedure. In the appendix, we present the SAS code of the NLP procedure [10, 11] used to find out the maximum likelihood estimates presented in our examples.

3. THE WEIBULL PARTICULAR CASE

Consider the general Weibull survival model obtained by considering (2.2) with (2.3). Assuming that the scale parameter, the shape parameter and the incidence probability are affected by covariate vectors $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$, respectively, let us to consider $p(\mathbf{z}_i)$ as a logit link, such as, $\log\Big(\frac{p(\mathbf{z}_i)}{1-p(\mathbf{z}_i)}\Big) = \eta_0 + \sum_{j=1}^{k} \eta_j z_{ij}$, the log-linear models $\log(\mu(\mathbf{x}_i)) = \alpha_0 + \sum_{j=1}^{k} \alpha_j x_{ij}$ and $\log(\gamma(\mathbf{y}_i)) = \beta_0 + \sum_{j=1}^{k} \beta_j y_{ij}$. Thus, the log-likelihood function for $\boldsymbol{\gamma}$, $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ is given by

$$
\begin{aligned}
l(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma} | \mathbf{x}, \mathbf{y}, \mathbf{z}) \;\propto\; & \sum_{i=1}^{n} \delta_i \Big[\mathbf{y}_i^t \boldsymbol{\beta} + e^{\mathbf{y}_i^t \boldsymbol{\beta}} \mathbf{x}_i^t \boldsymbol{\alpha} + e^{\mathbf{y}_i^t \boldsymbol{\beta}} \log(t_i) \Big] \\
& + \sum_{i=1}^{n} \delta_i \log(p(\mathbf{z}_i)) \;-\; \sum_{i=1}^{n} \delta_i (t_i \, e^{\mathbf{x}_i^t \boldsymbol{\alpha}})^{e^{\mathbf{y}_i^t \boldsymbol{\beta}}} \\
& + \sum_{i=1}^{n} (1-\delta_i) \log \Big[p(\mathbf{z}_i) + (1-p(\mathbf{z}_i)) \, e^{(-t_i \, e^{\mathbf{x}_i^t \boldsymbol{\alpha}})^{e^{\mathbf{y}_i^t \boldsymbol{\beta}}}} \Big] \,,
\end{aligned} \tag{3.1}
$$

where $p(\mathbf{z}_i)^{-1} = e^{-(\gamma_0 + \sum_{j=1}^{k} \gamma_j z_{ij})} (1 + e^{\gamma_0 + \sum_{j=1}^{k} \gamma_j z_{ij}})$, $\boldsymbol{\alpha}^t = (\alpha_0, ..., \alpha_k)$, $\boldsymbol{\beta}^t = (\beta_0, ..., \beta_k)$, $\boldsymbol{\gamma}^t = (\gamma_0, ..., \gamma_k)$, $\mathbf{x}_i^t = (1, x_{i1}, ..., x_{ik})$, $\mathbf{y}^t = (1, y_{i1}, ..., y_{ik})$ and $\mathbf{z}^t = (1, z_{i1}, ..., z_{ik})$.

4. SOME APPLICATIONS

4.1. A first application

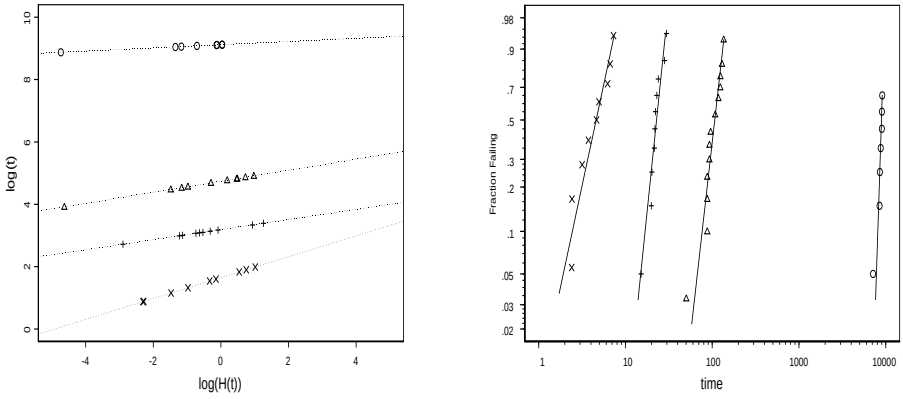
To check the assumption of shape parameter dependent on the covariates we can use graphical diagnostic methods. As a special case, consider the accelerated lifetime test (ALT) data on PET film, (see, Table 1), introduced by Hirose [12], see also Wang and Kececioglu ([30]). The ALT was performed at

Table 1: Failure times (hours) from an accelerated life test on PET film in SF_6 gas insulated transformers, [12].

Voltage	Failure times
5 kV	7131, 8482, 8559, 8762, 9026, 9034, 9104, 9104.25*, 9104.25*, 9104.25*
7 kV	50.25, 87.75, 87.76, 87.77, 92.90, 92.91, 95.96, 108.3, 108.3, 117.9, 123.9, 124.3, 129.7, 135.6, 135.6
10 kV	15.17, 19.87, 20.18, 21.50, 21.88, 22.23, 23.02, 28.17, 29.70
15 kV	2.40, 2.42, 3.17, 3.75, 4.65, 4.95, 6.23, 6.68, 7.30

Starred quantities denote censored observations.

four levels of the voltage; $v = 5, 7, 10$ and 15, with $10, 15, 10$ and 9 observations each, respectively. Three censored values were observed at $v = 5$. Denoting by $S(t) = P(T > t)$, the survival function, we should have parallel straight lines for the plots of $\log(-\log \hat{S}(t))$ versus $\log(t)$ for each stress level considering the Weibull distribution [16]. This is also true for the Weibull probability plot, Figure 1-b. In Figures 1-a and 1-b we observe straight lines which indicates that the Weibull distribution is appropriate, but we do not have parallel lines which indicates different shape parameters for each stress level. Interested readers can refer to Chapters 2, 7 and 8 of Meeker and Escobar ([22]), which present different methods to search for an appropriate lifetime distribution for fitting a set of data.



(a) (b)

Figure 1: Weibull fit for PET film data, Table 1.
(a): Hazard plot.
(b): Probability plot. 5 kV (○), 7 kV (△), 10 kV (+) and 15 kV (×).

Figure 1 indicate that the scale and shape parameter of the Weibull distribution should be affected by the stress levels. Moreover, following [30], plots show that $\log \hat{\mu}$ and $\log \hat{\gamma}$ have linear relationships with $x_1 = y_1 = -\log(v - 4.76)$,

where $\hat{\mu}$ and $\hat{\gamma}$ are the MLEs of μ and γ, obtained by considering each individual covariate level, which are given in Table 2, the constant 4.76 is a fixed threshold level [12], below which a failure is unlikely to occur.

Table 2: Maximum likelihood and standard deviation estimates considering a Weibull model for each stress level.

		MLE	
Level	log-likelihood	$\hat{\mu}$	$\hat{\gamma}$
5 kV	-57.7394	9.1145	2.9721
		(0.0196)	(0.3496)
7 kV	-67.5903	4.7367	1.7315
		(0.0480)	(0.2100)
10 KV	-28.1308	3.1873	1.8230
		(0.0541)	(0.2375)
15 kV	-17.4361	1.6474	1.0938
		(0.1179)	(0.2676)

Table 3 shows the MLEs for the parameter of (2.3) and their standard deviations assuming $\log(\mu(x_1)) = \alpha_0 + \alpha_1 \log(v - 4.76)$ and $\gamma(y_1) = \text{constant}$ (hereafter called Model A) and $\log(\gamma(y_1)) = \beta_0 + \beta_1 \log(v - 4.76)$ (hereafter called Model B).

Table 3: Maximum likelihood estimates considering two Weibull models.

		Estimates	
Model	Parameter	MLE	StDev
model A	α_0	6.3480	0.0399
	α_1	-1.9629	0.0265
	β	1.6080	0.1281
model B	α_0	6.3285	0.0213
	α_1	-1.9529	0.0156
	β_0	2.2311	0.1776
	β_1	-0.4636	0.1152

Locally at the MLEs, the values of the log-likelihood functions are -179.9849 (for Model A) and -173.2728 (for Model B). The values of the likelihood ratio statistics, Wald and score statistics to test model A against model B, that is, $H_0\colon \beta_1 = 0$ against $H_1\colon \beta_1 \neq 0$, are equal to 13.4240, 16.1896 and 17.0416, respectively. Their empirical p-values obtained from 10 000 bootstrap simulations are equal to 0.0007, 0.0007 and 0.0014, respectively, leading to a strong

evidence in favour of the complete model (Model B). The empirical distributions of these statistics are given in Figure 2 together with their Q-Q plots. We do not observe a good approximation to the chi-square distribution with one degree of freedom.

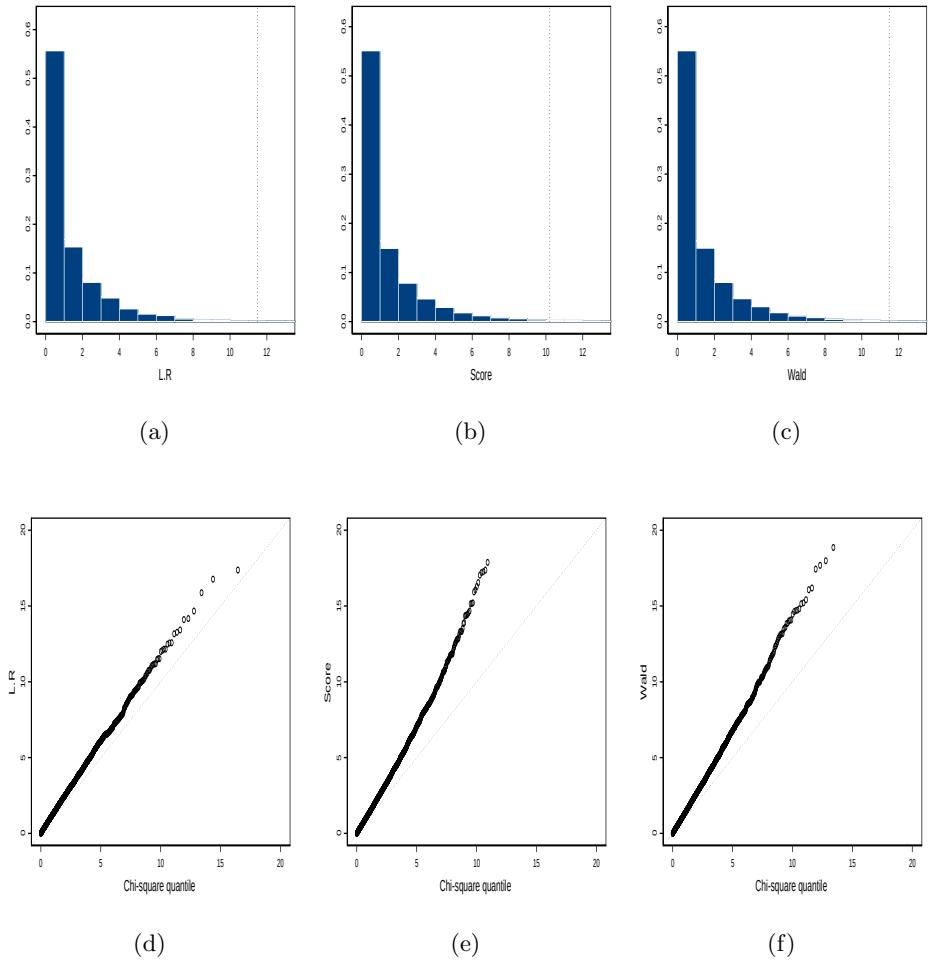


Figure 2: Empirical distributions. (a): Likelihood ratio statistic; (b): Score statistic; (c): Wald statistic; (d-f): Q-Q plots for (a), (b) and (c).

4.2. A second application

As an example where scale, shape and the proportion of immunes parameters may depend on covariates, consider the ovarian cancer data given by Edmunson *et al.* ([7]) and Therneau ([29]) (see also, [20] pp. 134 and [5] pp. 142).

The response variable (see, Table 4) was the survival time, in years, for 26 women following randomization to one or other of the two chemotherapy treatments. In Table 4 the censor indicator variable is 1 if t_i is an observed survival time and 0 if t_i is a right-censored observation. As pointed out in [20], we notice that large survival times tend to be censored, so there is some evidence of the existence of an immune component. To verify the possible difference between two treatments (treatment 1: standard chemotherapy — cyclophosphamide alone; treatment 2: combined chemotherapy — cyclophosphamide combined with adriamycin), [5] considered the usual Weibull regression model with covariates affecting only the scale parameter and concluded that there is a nonsignificant difference between the two treatments. In fact, the final model considered by Collett (1994) included age and treatment as covariates.

Table 4: Survival times (in years) of ovarian cancer patients.

Survival Time Group 1	Censor Indicator	Survival Time Group 2	Censor Indicator
0.1616	1	0.9671	1
0.3151	1	1.0000	1
0.4274	1	1.2712	1
0.7342	1	1.3014	1
0.9014	1	1.5425	1
1.1808	1	1.0329	0
1.7479	1	1.1534	0
1.2274	0	2.0384	0
1.3068	0	2.1068	0
2.2000	0	2.1096	0
2.3425	0	3.0932	0
2.8493	0	3.3041	0
3.0301	0	3.3616	0

From the survival curves (see, Figure 3), we observe that there are large censored observations, which could indicate the presence of immune individuals [20]. Therefore, we assume the model (2.2) with $S_0(t)$ given by (2.3) with $\log(\frac{p_i}{1-p_i}) = \eta_0 + \eta_1 x_i$, $\log(\mu_i) = \alpha_0 + \alpha_1 x_i$ and $\log(\gamma_i) = \beta_0 + \beta_1 x_i$, where x_i taking the value 1 if individual i is in the treatment group 1 or the value 2 if i is in the treatment group 2.

In this way, we can have the following hypothesis tests:

H_0: $\eta_1 = 0$ (no treatment effect in the proportion of cured patients),
H_0: $\alpha_1 = 0$ (no treatment effect in the ratio of susceptible patients) or
H_0: $\beta_1 = 0$ (no treatment effect in the shape of the lifetime distribution).

In Table 5 we have the MLE and their asymptotic standard-deviation estimates considering 4 models:

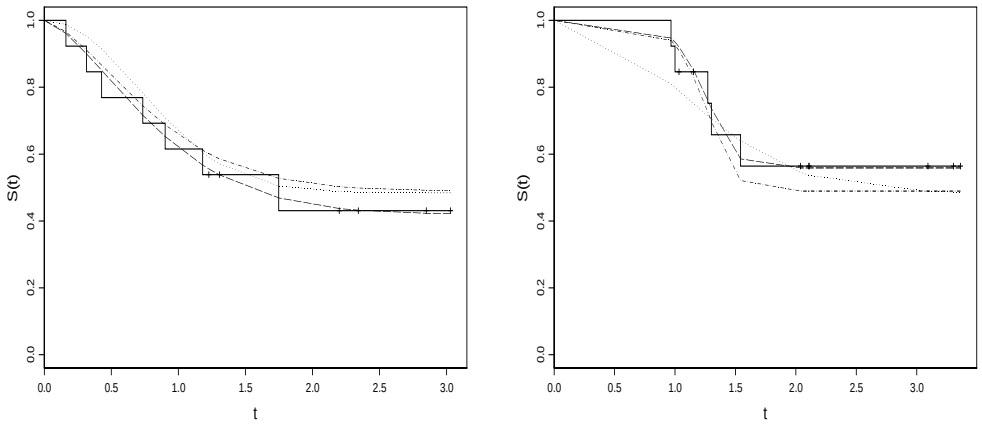
Model 1: $\log(\frac{p_i}{1-p_i}) = \eta_0,\ \log(\mu_i) = \alpha_0$ and $\log(\gamma_i) = \beta_0$;
Model 2: $\log(\frac{p_i}{1-p_i}) = \eta_0,\ \log(\mu_i) = \alpha_0 + \alpha_1 x_i$ and $\log(\gamma_i) = \beta_0$;
Model 3: $\log(\frac{p_i}{1-p_i}) = \eta_0,\ \log(\mu_i) = \alpha_0 + \alpha_1 x_i$ and $\log(\gamma_i) = \beta_0 + \beta_1 x_i$ and
Model 4: $\log(\frac{p_i}{1-p_i}) = \eta_0 + \eta_1 x_i,\ \log(\mu_i) = \alpha_0 + \alpha_1 x_i$ and $\log(\gamma_i) = \beta_0 + \beta_1 x_i$.

Locally at the MLE, the values of $-2\log(likelihood)$ are given by 49.3512 (Model 1), 48.1652 (Model 2), 40.6565 (Model 3) and 40.2318 (Model 4). We observe that Model 4 seems to give better fit for the data. This result is corroborated by Figure 3, where we have the plots of the fitted survival curves obtained from Models 2, 3 and 4 and the nonparametric Kaplan–Meier survival curve. We omitted the fitted survival curve from Model 1, which is very far from the Kaplan–Meier survival curve.

Table 5: Maximum likelihood estimates — long-term survivors models.

	Parameter					
Model	η_0	β_0	α_0	α_1	β_1	η_1
1	0.0284	0.7457	0.1423			
	(0.4300)	(0.2658)	(0.1572)			
2	0.0614	0.7222	-0.3759	0.3600		
	(0.4464)	(0.2663)	(0.5293)	(0.3764)		
3	0.0420	-1.0535	-0.4173	0.3482	1.4744	
	(0.4240)	(0.7314)	(0.5749)	(0.2936)	(0.4686)	
4	0.8870	-1.0782	-0.3627	0.3201	1.4833	-0.5614
	(1.3954)	(0.7615)	(0.6232)	(0.3175)	(0.4812)	(0.8725)

It is important to point out that in this application we have a small data set (26 patients) and should be careful to conclude that model 4 provides a better fit. In fact, model 3 and model 4 give similar fits for the survival curves (see Figure 3) and the difference $40.6565 - 40.2318 = 0.4247$ is nonsignificant. In this case the cured proportions and rates of failure do not seem to differ significantly between the treatment groups.



(a) (b)

Figure 3: Survival curves.
(a): standard chemotherapy;
(b): combined chemotherapy;
(− − −) Kaplan–Meier;
(· · ·) model 2;
(— · —) model 3;
(— —) model 4.

5. CONCLUDING REMARKS

In this paper, we consider a general class of survival models where the shape, the scale and the incidence probability parameters can be dependent on covariates. The major advantage of the general survival class of models lies on its ability to accommodate several usual survival models. From the practical viewpoint the methodology can be implemented straightforwardly and runs immediately using existing statistical packages.

ACKNOWLEDGMENTS

This work has been supported by the Brazilian Organizations CAPES and CNPq. The authors would like to thank the referees for their helpful comments.

Appendix A — Maximum Likelihood, First Application

In this appendix, we present the SAS code used to get the maximum likelihood estimates presented the both examples. The optimization of the log-likelihood was made by using the nonlinear programming SAS procedure considering the trust-region algorithm, [6, 24].

Listing 1: Single Weibull model.

```
proc nlp data=hirose tech=tr phes cov=2 vardef=n;
 max L;
 parms alpha0 = 6.0, beta0  = 1.0;
 mu     = exp(alpha0);
 beta   = exp(beta0);
 logH   = log(beta)-beta*log(mu)+beta*log(t);
 logS   = -(t/mu)**beta;
 L      = delta*logH+logS;
 by voltage;
run;
```

Listing 2: Weibull model with constant shape parameter.

```
proc nlp data=hirose tech=tr phes cov=2 vardef=n;
 max L;
 parms alpha0 = 6.0, alpha1 = 0.9, beta0 = 1.0;
 mu     = exp(alpha0+alpha1*voltage);
 beta   = exp(beta0);
 logH   = log(beta)-beta*log(mu)+beta*log(t);
 logS   = -(t/mu)**beta;
 L      = delta*logH+logS;
run;
```

Listing 3: Weibull model with nonconstant shape parameter.

```
proc nlp data=hirose tech=tr phes cov=2 vardef=n;
 max L;
 parms
 alpha0 = 6.0, alpha1 = -2.0, beta0 = 2.2, beta1 = -0.4;
 mu     = exp(alpha0+alpha1*voltage);
 beta   = exp(beta0+beta1*voltage);
 logH   = log(beta)-beta*log(mu)+beta*log(t);
 logS   = -(t/mu)**beta;
 L      = delta*logH+logS;
run;
```

Appendix B — Maximum Likelihood, Second Application

Listing 4: Long-term survivors model — model 4.

```
proc nlp data = dados tech=tr cov=2 vardef=n phes;
  max L;
  parms  alpha0 =-0.4, alpha1 = 0.3, beta0  = -1.0,
         beta1  = 1.4, g0     = 0.0, g1     = 0.0;
  mu     = exp(alpha0+alpha1*treatment);
  beta   = exp(beta0+beta1*treatment);
  p      = exp(g0+g1*x1)/(1+exp(g0+g1*treatment));
  h      = (beta/mu)*(t/mu)**(beta-1);
  S      = exp(-(t/mu)**beta);
  Lc     = log(p)+log(h)+log(S);
  Li     = log(1-p+p*S);
  L      = delta*Lc+(1-delta)*Li;
run;
```

REFERENCES

[1] Anderson, M.K. (1991). A nonproportional hazards Weibull accelerated failure time regression model, *Biometrics*, **47**, 281–288.

[2] Barbosa, E.P.; Colosimo, E.A., and Louzada-Neto, F. (1996). Accelerated life tests analysed by a piecewise exponential distribution via generalized linear models, *IEEE Transactions on Reliability*, **45**, 4, 619–623.

[3] Cepeda, E., and Gamerman, D. (2000). Bayesian modeling of variance heterogeneity in normal regression models, *Brazilian Journal of Probability and Statistics*, **14**, 207–221.

[4] Chan, C.K. (1991). Temperature-dependent standard deviation of log(failure time) distributions, *IEEE Transactions on Reliability*, **40**, 2, 157–160.

[5] Collett, D. (1994). *Modelling Survival Data in Medical Research*, Chapman and Hall, New York.

[6] Dennis, J.E.; Gay, D.M. and Welsch, R.E. (1981). An adaptive nonlinear least-squares algorithm, *ACM Transactions on Mathematical Software*, **7**, 348–368.

[7] Edmunson. J.H.; Fleming, T.R.; Decker, D.G.; Malkasian, G.D.; Jorgenson, E.O.; Jeffries, J.A.; Webb, M.J. and Kvols, L.K. (1979). Different chemotherapeutic sensitivities and host factors affecting prognosis in advanced ovarian carcinoma versus minimal residual disease. *Cancer Treatment Reports*, **63**, 241–247.

[8] Farewell, V.T. (1982). The use of mixture models for the analysis of survival data with long-term survivors, *Biometrics*, **38**, 1041-1046.

[9] Green, P.J. (1984). Iteratively re-weighted least squares for maximum likelihood estimation and some robust alternatives, *Journal of the Royal Statistical Society B*, **46**, 149–192.

[10] Hartmann, W. (1992). *Applications of Nonlinear Optimization with PROC NLP and SAS/IML Software*, Technical Report, Cary, N.C.: SAS Institute Inc.

[11] Hartmann, W. (1992). *The NLP Procedure: Extended User's Guide*, Cary: SAS Institute Inc.

[12] Hirose, H. (1993). Estimation of threshold stress in accelerated life-testing, *IEEE Transactions on Reliability*, **42**, 650–657.

[13] Hsie, F. (2001). On heteroscedastic hazards regression models: theory and application, *Journal of the Royal Statistical Society B*, **63**, 1, 63–79.

[14] Ihaka, R. and Gentleman, R.R. (1996). A language for data analysis and graphics. *Journal of Computational and Graphical Statistics*, **5**, 299–314.

[15] Jensen, S.T.; Johansen, S. and Lauritzen, S.L. (1991). Globally convergent algorithm for maximizing a likelihood function, *Biometrika*, **78**, 4, 867–877.

[16] Lawless, J.F. (1982). *Statistical Models and Methods for Lifetime Data*, John Wiley, New York.

[17] Louzada-Neto, F. (1997). Extended hazard regression model for reliability and survival analysis, *Lifetime Data Analysis*, **3**, 367–381.

[18] Louzada-Neto, F. (1999). Modelling lifetime data by hazard models: a graphical approachs, *Applied Stochastic Models in Business and Industry*, **15**, 123–129.

[19] Louzada-Neto, F. (2001). Bayesian analysis for hazard models with nonconstant shape parameters, *Computational Statistics*, **16**, 243-254.

[20] Maller, R.A. and Zhou, X. (1996). *Survival Analysis with Long-Term Survivors*, John Wiley, New York.

[21] Meeker, W.Q. and Escobar, L.A. (1993). A review of recent research and current issues in accelerated testing, *International Statistical Review*, **61**, 1, 147–168.

[22] Meeker, W.Q. and Escobar, L.A. (1998). *Statistical Methods for Reliability Data*, John Wiley, New York.

[23] Meeter, C.A. and Meeker, W.Q. (1994). Optimum accelerated life tests with a nonconstant scale parameter, *Technometrics*, **36**, 71–83.

[24] Moré, J.J. and Sorensen, D.C. (1983). Computing a Trust-Region Step, *SIAM Journal on Scientific and Statistical Computing*, **4**, 553–572.

[25] Nelson, W. (1984). Fitting of fatigue curves with nonconstant standard deviation to data with runouts, *Journal of Testing and Evaluation*, **12**, 69–77.

[26] Pascual, F.G. and Meeker, W.Q. (1997). Regression analysis of fatigue data with runouts based on a model with nonconstant standard deviation and a fatigue limit parameter, *Journal of Testing and Evaluation*, **25**, 292–301.

[27] Smith, R.L. (1991). *Weibull regression models for reliability data*. In "Reliability Engineering and System Safety", Elsevier Publishers, 55–77.

[28] Tanizaki, H. and Zhang, X. (2001). Posterior analysis of the multiplicative heterocedasticity model, *Communications in Statistics. Theory and Methods*, **30**, 5, 855–874.

[29] Therneau, T.M. (1986). *The COXREGR Procedure*, In "SAS SUGI Supplemental Library Users's Guide, Version 5", SAS Institute Inc., Gary, IN.

[30] Wang, W. and Kececioglu, D.B. Fitting the Weibull log-linear model to accelerated life-test data, *IEEE Transactions on Reliability*, **49**, 2, 217–223.

[31] Zhou, X., Stroupe, K.T. and Tierney, W.M. (2001). Regression analysis of health care charges with heteroscedasticity, *Applied Statistics*, **50**, 3, 303–312.