UNIVERSIDADE DE SÃO PAULO
Instituto de Ciências Matemáticas e de Computação
ISSN 0103-2569

Bayesian Inference for PolyHazard Models in the Presence of Covariates

Josmar Mazucheli
Francisco Louzada-Neto
Jorge Alberto Achcar

Nº 119

RELATÓRIOS TÉCNICOS

São Carlos – SP
Ago./2000

Bayesian Inference for PolyHazard Models in the Presence of Covariates

Josmar Mazucheli*
Universidade Estadual de Maringá - DEs/UEM,
C.P. 870, 87020-900, Maringá, P.R. - Brazil.

Francisco Louzada-Neto
Universidade Federal de São Carlos - DEs/UFSCar,
C.P. 676, 13565-905, São Carlos, S.P. - Brazil.

Jorge Alberto Achcar
Universidade de São Paulo - ICMC/USP,
C.P. 668, 13560-970, São Carlos, S.P. - Brazil

Abstract

The polyhazard models are a flexible family for fitting lifetime data. Their main advantage over the single hazard models, such as the Weibull and the log-logistic models, is to acommodate a large amount of non-monotone hazard shapes, including bathtub and multimodal curves. The main goal of this paper is to present a Bayesian inference procedure for the polyhazard models in the presence of covariates, generalizing the Bayesian analysis presented in Berger and Sun (1993), Basu, *et. al.* (1999) and Kuo and Yang (2000). The two most important particular polyhazard models, namely poly-Weibull, poly-log-logistc and a combination of both are studied in detail. The methodology is illustrated in two real medical datasets.

Key words and phrases: Bayesian inference, polyhazard models, Gibbs Sampling algorithm, Arms algorithm, artificial variables, Kaplan-Meier survival curves, Bayes factor.

*Josmar Mazucheli is a graduate student at COPPE/UFRJ-Brazil.

1 Introduction

Let us assume the situation where a unit can have $k \geq 2$ possible causes of failure, such that the exact cause is fully or partial unknow (see for example Basu, *et. al.*, 1999; Kuo and Yang, 2000). If $X_j, j = 1, \ldots, k$ denote the time to failure due to the j^{th} cause, then the observed random variable is $T = \min(X_1, \ldots, X_k)$ which is said to have a polyhazard distribution (see Louzada-Neto, 1999) with hazard function given by

$$h(t) = \sum_{j=1}^{k} h_j(t). \tag{1}$$

A special case of (1) is given by the poly-Weibull with hazard function, (Davison and Louzada-Neto, 1999)

$$h(t) = \sum_{j=1}^{k} \frac{\beta_j t^{\beta_j - 1}}{\mu_j}, \tag{2}$$

where $\mu_j > 0$ and $\beta_j > 0$ are respectively the scale and shape parameters associated to each component. These k-Weibull hazard functions can have different scale parameters, allowing wide flexibility into the model.

The polyhazard distributions are commonly used for competing risks problems considering either, machine or biological systems, where a cause of failure may be a machine-component or a certain disease, and a failure may be a non functional state or death. The polyhazard distributions give greater flexibility to fit lifetime data since the hazard function supports a rich class of hazard shapes. For instance, for $k = 2$ in (2), the hazard is increasing if $\min(\beta_1, \beta_2) > 1$, decreasing for $\max(\beta_1, \beta_2) < 1$ and bathtub shaped for $\beta_1 < 1$ and $\beta_2 > 1$ (Berger and Sun, 1993).

Inferences for the poly-Weibull distributions are introduced by many authors. For example, Davison and Louzada-Neto (1999), Berger and Sun (1993, 1996), Basu *et. al.* (1999).

Other choices for the hazard components in (1) could be considered in applications. A special case is given by the poly-log-logistic distribution with hazard function

$$h(t) = \sum_{j=1}^{k} \frac{\beta_j t^{\beta_j - 1}}{\mu_j + t^{\beta_j}}, \tag{3}$$

or by a combination of (2) and (3), called poly-Weibull-log-logistic hazard model, given by

$$h(t) = \sum_{j=1}^{d} \frac{\beta_j t^{\beta_j - 1}}{\mu_j} + \sum_{j=d+1}^{k} \frac{\beta_j t^{\beta_j - 1}}{\mu_j + t^{\beta_j}}. \tag{4}$$

Usually there are situations where the failure time may depend on a vector $\mathbf{z}$ of explanatory variables. It therefore becomes of interest consider generalizations of (1) including special functions of $\mathbf{z}$ affecting each hazard component. The polyhazard models can be extended to include regressor variables in different ways. However, the most commonly used are given by,

$$h(t) = \sum_{j=1}^{k} \frac{\beta_j t^{\beta_j - 1}}{\mu_j} e^{\mathbf{z}^t \boldsymbol{\alpha}_j}, \tag{5}$$

for the poly-Weibull, corresponding to assume proportionality between the hazards (Cox, 1972), and

$$h(t) = \sum_{j=1}^{k} \frac{\beta_j t^{\beta_j - 1} e^{\mathbf{z}^t \boldsymbol{\alpha}_j}}{\mu_j + t^{\beta_j} e^{\mathbf{z}^t \boldsymbol{\alpha}_j}}, \tag{6}$$

for the poly-log-logistic, where $\boldsymbol{\alpha}_j$ is the vector of regression parameters. Extensions could also be considered allowing the shape parameters to be depend on covariates but it will not be considered here.

The main goal of this paper is to present a Bayesian analysis of polyhazard models in the presence of covariates and show that these models can provide better fit than single models such as the single Weibull and the single log-logistic. In this way, our study generalizes the Bayesian analysis considered in Berger and Sun (1993), Basu, *et. al.* 1999 or yet Kuo and Yang (2000) assuming that the failure dependents on covariates. Sections 2 and 3 describe a Bayesian formulation for the poly models. Numerical examples considering two real data sets are presented in Section 4. Some concluding remarks in Section 5 finalize the paper.

2 The Poly-Weibull Distribution

Consider a sample of independent random variables $T_1, \ldots, T_n$ such that $T_i = \min(X_{i1}, \ldots, X_{ik})$ and T_i has an associated covariate vector $\mathbf{z}_i^t = (z_{i1}, \ldots, z_{ip})$ and an indicator variable defined by $\delta_i = 1$ if t_i is an observed failure time and $\delta_i = 0$ if t_i is a right-censored observation. In this way, the likelihood function for the parameters of any set of survival data subject to uninformative censoring can be written as,

$$L(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \prod_{i=1}^{n} h(t_i)^{\delta_i} S(t_i). \tag{7}$$

Assuming the poly-Weibull (5) and the reparametrization $\alpha_{0j} = -\log(\mu_j)$, the survival function for t_i is given by,

$$S(t_i) = \exp\left[-\sum_{j=1}^{k} t_i^{\beta_j} e^{\mathbf{z}_i^t \boldsymbol{\alpha}_j}\right], \tag{8}$$

where $\mathbf{z}_i^t\boldsymbol{\alpha}_j = \alpha_{0j} + \alpha_{1j}z_{i1} + \ldots + \alpha_{pj}z_{ip}$, $\boldsymbol{\alpha}^t = (\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_k)$, $\boldsymbol{\alpha}_j = (\alpha_{0j}, \ldots, \alpha_{pj})$ and $\boldsymbol{\beta}^t = (\beta_1, \ldots, \beta_k)$.

Let us assume a prior with independent components for $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, $\pi(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \pi(\boldsymbol{\alpha})\pi(\boldsymbol{\beta})$. From (7), (5) and (8), the joint posterior distribution for $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ is given by,

$$\pi(\boldsymbol{\alpha}, \boldsymbol{\beta}|\mathbf{t}, \mathbf{z}) \propto \pi(\boldsymbol{\alpha})\pi(\boldsymbol{\beta}) \prod_{i=1}^{n} \left(\sum_{j=1}^{k} \beta_j t_i^{\beta_j - 1} e^{\mathbf{z}_i^t \boldsymbol{\alpha}_j} \right)^{\delta_i} \exp\left(- \sum_{i=1}^{n} \sum_{j=1}^{k} t_i^{\beta_j} e^{\mathbf{z}_i^t \boldsymbol{\alpha}_j} \right). \quad (9)$$

In (9) the term $\prod_{i=1}^{n} \left(\sum_{j=1}^{k} \beta_j t_i^{\beta_j - 1} e^{\mathbf{z}_i^t \boldsymbol{\alpha}_j} \right)^{\delta_i}$ prevents us to get the conditional densities as a product of independent components. To simplify the joint posterior and the full conditional distributions for the Gibbs Sampling algorithm we introduce artificial variables (see for example Tanner and Wong, 1987), $\mathbf{v}_i = (v_{i1}, \ldots, v_{ik})$, for each uncensored observation and assume that the conditional distribution for $\mathbf{v}_i$ given $\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{t}$ and $\mathbf{z}$ has a multinomial distribution, $\text{Mult}(1; u_{i1}, \ldots, u_{ik})$, with cell probabilities,

$$u_{ij} = \frac{h_j(t_i)}{\sum_{j=1}^{k} h_j(t_i)}, \quad (10)$$

that is, for each uncensored observation,

$$\pi(\mathbf{v}_1, \ldots, \mathbf{v}_n|\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{t}, \mathbf{z}) \propto \frac{\prod_{i=1}^{n} \prod_{j=1}^{k} h_j(t_i)^{\delta_i v_{ij}}}{\prod_{i=1}^{n} \left(\sum_{j=1}^{k} h_j(t_i) \right)^{\delta_i}}, \quad (11)$$

where $h_j(t_i) = \beta_j t_i^{\beta_j - 1} e^{\mathbf{z}_i^t \boldsymbol{\alpha}_j}$, for $j = 1, \ldots, k$, $i = 1, \ldots, n$ and $\sum_{j=1}^{k} v_{ij} = 1$.

Combining equation (11) with (9) we have the joint posterior distribution, given $\mathbf{t}, \mathbf{z}$ and $\mathbf{v} = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ as a product of independent components,

$$\pi(\boldsymbol{\alpha}, \boldsymbol{\beta}|\mathbf{t}, \mathbf{z}, \mathbf{v}) \propto \pi(\boldsymbol{\alpha})\pi(\boldsymbol{\beta}) \prod_{i=1}^{n} \prod_{j=1}^{k} \left(\beta_j t_i^{\beta_j} e^{\mathbf{z}_i^t \boldsymbol{\alpha}_j} \right)^{\delta_i v_{ij}} \exp\left(- \sum_{i=1}^{n} \sum_{j=1}^{k} t_i^{\beta_j} e^{\mathbf{z}_i^t \boldsymbol{\alpha}_j} \right). \quad (12)$$

To generate samples of the joint posterior distribution (12) we follow the steps:

i) Start with initial values $\boldsymbol{\alpha}^{(0)}$ and $\boldsymbol{\beta}^{(0)}$;

ii) generate $\mathbf{v}^{(1)} = \left(v_1^{(1)}, \ldots, v_n^{(1)} \right)$ from a multinomial distribution with cell probabilities (10);

iii) generate a sample of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ from the conditional distributions, $\pi\left(\boldsymbol{\alpha}|\boldsymbol{\beta}^{(0)}, \mathbf{v}^{(1)}, \mathbf{t}, \mathbf{z}\right)$ and $\pi\left(\boldsymbol{\beta}|\boldsymbol{\alpha}^{(1)}, \mathbf{v}^{(1)}, \mathbf{t}, \mathbf{z}\right)$.

Assuming prior independence among the parameters we consider the following prior densities for $\log(\beta_j)$ and α_{lj}, $j = 1, \ldots, k$, $l = 0, \ldots, p$,

$$\beta_j \sim N(a_j, b_j), \alpha_{lj} \sim N(c_{lj}, d_{lj}), \tag{13}$$

where a_j, b_j, c_{lj}, d_{lj} are known.

We note however that usual non informative prior can lead to improper posterior (Davison and Louzada-Neto, 1999).

Thus, the conditional posterior distributions for the sampling schemes are given by,

$$\pi\left(\beta_j | \boldsymbol{\beta}_{(j)}, \boldsymbol{\alpha}, \mathbf{t}, \mathbf{z}, \mathbf{v}\right) \propto \beta_j^{n_j} \exp\left(\beta_j \sum_{i=1}^{n} \delta_i v_{ij} \ln(t_i) - \sum_{i=1}^{n} t_i^{\beta_j} e^{\mathbf{z}_i^t \alpha_{jl}} - \frac{1}{2}\left(\frac{\beta_j - a_j}{b_j}\right)^2\right), \tag{14}$$

where $n_j = \sum_{i=1}^{n} \delta_i z_{ij}$, $\boldsymbol{\beta}_{(j)} = (\beta_1, \ldots, \beta_{j-1}, \beta_{j+1}, \ldots, \beta_k)$ and

$$p\left(\alpha_{jl} | \boldsymbol{\alpha}_{(\alpha_{jl})}, \boldsymbol{\beta}, \mathbf{t}, \mathbf{z}, \mathbf{v}\right) \propto \exp\left(\sum_{i=1}^{n} \delta_i v_{ij} \mathbf{z}_i \alpha_{jl} - \sum_{i=1}^{n} t_i^{\beta_j} e^{\mathbf{z}_i^t \alpha_{jl}} - \frac{1}{2}\left(\frac{\alpha_{jl} - c_{jl}}{d_{jl}}\right)^2\right), \tag{15}$$

where $l = 0, 1, \ldots, p$, $j = 1, \ldots, k$, $z_{i0} = 1$ and $\boldsymbol{\alpha}_{(\alpha_{jl})} = (\alpha_{01}, \ldots, \alpha_{j-1,l-1}, \alpha_{j+1,l+1}, \ldots, \alpha_{pk})$.

From the conditional posterior densities for poly-Weibull distribution we can see that standard sampling schemes are not feasible since the conditional distributions are not of a commom form. Bayesian inference for the parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ can be however performed by Metropolis-Hastings algorithms (see for example Chib and Greenberger, 1995) considering the conditionals as the target densities. In our applications we have used the Adaptive Rejection Metropolis Sampling algorithm (arms) introduced by Gilks, *et. al.* (1995) and also discussed in Gilks (1996), Gilks, *et. al.* (1997). The arms algorithm is a generalization of the method of adaptive rejection sampling of Gilks (1992), Gilks and Wild (1992) which includes a Metropolis step to accomodate non-log-concavity in the density that will be sampled. The C code written by Gilks and linked to the matrix language Ox (see Doornik, 1999) was used to get the posterior summaries of interest. In our examples, arms works better than Metropolis-Hastings algorithms since the suitable candidate-generating density are not obvious to be specified. However, others schemes as accept-reject for the log-concave densities or by using auxiliary variables (Damien, *et. al.* 1999) could be considered in principle. Alternatives to Bayesian estimation such as maximum likelihood estimates (MLEs) can be obtained straightforward by direct maximization of the likelihood (7) or using EM algorithm (Dempster *et. al.* 1977).

3 The Poly-Log-Logistic-Distribution

Assuming the poly-log-logistic model (6) with the reparametrization $\alpha_{0j} = -\log(\mu_j)$, the survival function for t_i is given by $S(t_i) = \left(1 + t_i^{\beta_j} e^{\mathbf{z}_i^t \boldsymbol{\alpha}_j}\right)^{-1}$, where $\mathbf{z}_i^t \boldsymbol{\alpha}_j = \alpha_{0j} + \alpha_{1j} z_{i1} + \ldots + \alpha_{pj} z_{ip}$.

It follows that the likelihood function for $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ assuming right-censored observations is given by,

$$L(\boldsymbol{\alpha}, \boldsymbol{\beta}) \propto \prod_{i=1}^{n} \left(\sum_{j=1}^{k} \frac{\beta_j t_i^{\beta_j} e^{\mathbf{z}_i^t \boldsymbol{\alpha}_j}}{1 + t_i^{\beta_j} e^{\mathbf{z}_i^t \boldsymbol{\alpha}_j}} \right)^{\delta_i} \prod_{j=1}^{k} \left(1 + t_i^{\beta_j} e^{\mathbf{z}_i^t \boldsymbol{\alpha}_j}\right)^{-1}. \quad (16)$$

Let us assume the same prior distributions (13) and also consider the introduction of the latent variable $\mathbf{v}_i = (v_{i1}, \ldots, v_{ik})$, where v_i has a multinomial distribution with u_{ij} defined in (10), such as,

$$h_j = \frac{\beta_j t_i^{\beta_j - 1} e^{\mathbf{z}_i^t \boldsymbol{\alpha}_j}}{1 + t_i^{\beta_j} e^{\mathbf{z}_i^t \boldsymbol{\alpha}_j}}. \quad (17)$$

Thus, the joint posterior distribution for $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, given $\mathbf{t}, \mathbf{z}$ and $\mathbf{v} = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$, is given by,

$$\pi(\boldsymbol{\alpha}, \boldsymbol{\beta} | \mathbf{t}, \mathbf{z}, \mathbf{v}) \propto \pi(\boldsymbol{\alpha})\, \pi(\boldsymbol{\beta}) \prod_{i=1}^{n} \prod_{j=1}^{k} \left(\frac{\beta_j t_i^{\beta_j} e^{\mathbf{z}_i^t \boldsymbol{\alpha}_j}}{1 + t_i^{\beta_j} e^{\mathbf{z}_i^t \boldsymbol{\alpha}_j}} \right)^{\delta_i v_{ij}} \prod_{j=1}^{k} \left(1 + t_i^{\beta_j} e^{\mathbf{z}_i^t \boldsymbol{\alpha}_j}\right)^{-1}, \quad (18)$$

and the conditional distributions are given by,

$$\begin{aligned} \pi\left(\beta_j | \boldsymbol{\beta}_{(j)}, \boldsymbol{\alpha}, \mathbf{t}, \mathbf{z}, \mathbf{v}\right) \quad \propto \quad & \beta_j^{n_j} \exp\left[\sum_{i=1}^{n} \delta_i z_{ij} \ln t_i^{\beta_j} \left(1 + t_i^{\beta_j} e^{\mathbf{z}_i^t \alpha_{jl}}\right)^{-1}\right] \quad (19) \\ & \exp\left[-\frac{1}{2}\left(\frac{\beta_j - a_j}{b_j}\right)^2\right] \prod_{i=1}^{n} \left(1 + t_i^{\beta_j} e^{\mathbf{z}_i^t \alpha_{jl}}\right)^{-1}, \end{aligned}$$

$$\begin{aligned} p\left(\alpha_{jl} | \boldsymbol{\alpha}_{(jl)}, \boldsymbol{\beta}, \mathbf{t}, \mathbf{z}, \mathbf{v}\right) \quad \propto \quad & \exp\left[\sum_{i=1}^{n} \delta_i z_{ij} \ln \frac{\alpha_{jl}}{\left(1 + t_i^{\beta_j} e^{\mathbf{z}_i^t \alpha_{jl}}\right)}\right] \quad (20) \\ & \exp\left[-\frac{1}{2}\left(\frac{\alpha_{jl} - c_{jl}}{d_{jl}}\right)^2\right] \prod_{i=1}^{n} \left(1 + t_i^{\beta_j} e^{\mathbf{z}_i^t \alpha_{jl}}\right)^{-1}. \end{aligned}$$

Bayesian inference for $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ can be obtained similarly to the scheme adopted for the poly-Weibull in Section 2.

4 Some Examples

4.1 Laryngeal Cancer Data

Consider the data from Table D.4 of Klein and Moeschberger (1997, p. 475) on the times (in years) between the first treatment and either death or the end of the study of 90 males diagnosed with cancer of the larynx during the period 1970–1978 at a Dutch Hospital.

Interest is centred here in discovering if prognostic effect of the stage of the patient's cancer affects significantly the patient's survival.

After preliminary investigations, it was discovered that a strong prognostic effect of the stage of the patient's cancer is indicated. The stages are based on the type of tumor, nodal involvement and distant metastasis grading, which are used by the American Joint Committee for Cancer Staging. The models presented in this paper are being illustrated on the times to death of 90 males related to the cancer stages; 18 deaths and 15 censoring in stage I, 7 death and 10 censoring in stage II, 17 death and 10 censoring in stage III and 11 death and 3 censoring for stage IV.

In Figure 1 we have the Kaplan-Meier estimates of the survival at each stages and the fits of survival curves considering the single Weibull, single log-logistic and poly models (poly-Weibull, poly-log-logistic and poly-Weibull-log-logistic) with a covariate for each stage. These results suggest that the poly models with a covariate are better than the single ones with a covariate.

The parameters for all models were estimated using the adaptive rejection Metropolis sampling algorithm where we simulated two separate Gibbs chains, each rans for 105000 iterations. In order to diminish the effect of the starting parameter values we discarded the first 5000 elements of each chain. Convergence of the Gibbs algorithm was observed using diagnostics procedures available in CODA package (Best, *et. al.* 1995). For each parameter we considered every 10^{th} draw and so we finally got a sample of size 20000.

The Monte Carlo estimates, in log-scale, for the posterior means of the parameters obtained from the combined chains are given in Table 1 for the single Weibull and single log-logistic and in Table 2 for the poly models. The hyperparameters, in all fitting, are setting such that we have proper but very weak prior.

Table 1: Posterior Mean and 95% Credible Interval.

	Model	
Parameter	Single Weibull	Single log-logistic
β_1	0.0790 (-0.1641;0.3056)	0.3108 (0.1075;0.5021)
α_{01}	2.1856 (1.6985;2.7174)	2.1949 (1.7354;2.6892)
α_{11}	0.5232 (0.2469;0.7981)	0.7537 (0.4430;1.0734)

Table 2: Posterior Mean and 95% Credible Interval.

Parameter	Model		
	Poly-Weibull	Poly-log-logistic	Poly-Weibull-log-logistic
β_1	0.3248	0.4853	0.3655
	(0.0198;0.6051)	(0.1676;0.7771)	(0.0757;0.6371)
α_{01}	3.3883	3.5231	3.4915
	(2.6953;4.1927)	(2.7009;4.466)	(2.7151;4.3645)
α_{11}	-0.0899	-0.0472	-0.0981
	(-0.7402;0.5047)	(-0.9008;0.6858)	(-0.7583;0.5046)
β_2	-0.1418	0.0127	0.0176
	(-0.6596;0.2166)	(-0.5601;0.3902)	(-0.5021;0.3689)
α_{12}	1.4150	1.6866	1.703
	(0.6777;2.0581)	(0.7637;2.5808)	(0.786;2.5548)

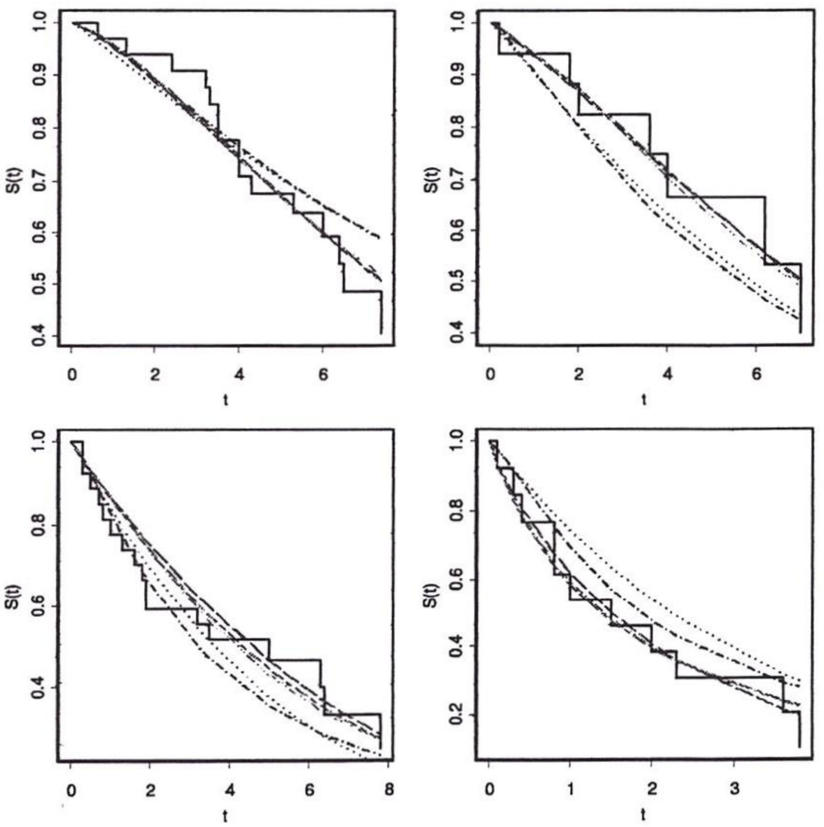

Figure 1: Fit of a single-models and a poly-models for each stage.(_) Kaplan-Meier survival curve, (...) single-Weibull, (._.) single-log-logistic, (— —) poly-weibull, (....) poly-log-logistic and (- — -) poly-weibull-log-logistic fits.

The Bayes factor could also be used to decide for the best model to be fitted. The Bayes factor (see for example Gelfand, 1996) is the relative weight of evidence for model M_1 compared to model M_2 given by $B_{12} = f(t_{obs}|M_1)/f(t_{obs}|M_2)$, where t_{obs} denotes the actual observations and $f(t_{obs}|M_k)$ denotes the marginal density under model $M_k, k = 1, 2$. In general, it can be useful to consider twice the logarithm of the Bayes Factor, which is on the same scale as the deviance and the likelihood ratio test statistics. Interpretation for this quantity was introduced by Jeffreys (1961).

Considering the Monte Carlo estimates for the marginal densities based on the generated Gibbs samples for the joint posterior distribution for the parameters, we have in Table 3 the values for $2\log(B_{lk}); l, k = 1, \ldots, 5,\ l \neq k$, where M_1 denotes de single Weibull model; M_2 denotes the single log-logistic model; M_3 denotes de poly-Weibull model; M_4 denotes de poly-log-logistic model and M_5 denotes the poly-Weibull-log-logistic model. We observe that models M_3, M_4 and M_5 give better fit for the laryngeal cancer data. In particular, model M_5 (poly-Weibull-log-logistic) is the best for the data.

Table 3: Bayes Factors B_{lk}.

$B_{21} = 1.6044$	$B_{42} = 1.7408$
$B_{31} = 3.6198$	$B_{52} = 2.3310$
$B_{41} = 3.3452$	$B_{43} = -0.2746$
$B_{51} = 3.9354$	$B_{53} = 0.3156$
$B_{32} = 2.0154$	$B_{54} = 0.5902$

4.2 Interstitial Cell Tummours Data

In a carcinogenesis experiment (National Toxicology Program, 1986), 100 male F344 rats were exposed by gavage to two dose levels of Commercial Grad toluene diisocyanate (Lagakos and Louis, 1988). Half of then where in the control group and half received 60 mg/kg of the component.

Figure 2 shows the Kaplan-Meier survival curves with fits for the single Weibull, single log-logistic and poly models. We observe close agreement between the Kaplan-Meier survival curves with the polyhazards models. The use of single Weibull or log-logistic models give poor fit for the data. The Monte Carlo estimates are showed in Tables 4 and 5 and obtained in a similar way as to the laryngeal cancer data.

The graphical results are corroborated by the Bayes Factor results in Table 6.

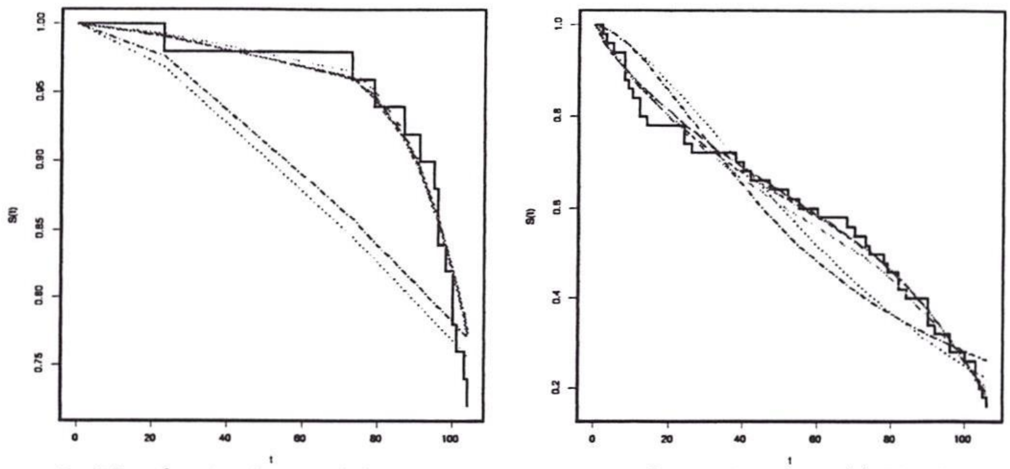

Figure 2: Fit of a single-models and a poly-models for each stage.(_) Kaplan-Meier survival curve, (...) single-Weibull, (._.) single-log-logistic, (__ __) poly-Weibull, (....) poly-log-logistic and (_ __ _) poly-Weibull-log-logistic fits.

Table 4: Posterior Mean and 95% Credible Interval.

Parameter	Model	
	Single Weibull	Single log-logistic
β_1	0.3772	0.5327
	(0.1397;0.6108)	(0.2923;0.7551)
α_{01}	7.278	8.0770
	(5.7904;9.0089)	(6.3967;9.9184)
α_{11}	1.6847	2.2496
	(1.1092.3109)	(1.474;3.0705)

Table 5: Posterior Mean and 95% Credible Interval.

Parameter	Model		
	Poly-Weibull	Poly-log-logistic	Poly-Weibull-log-logistic
β_1	2.0137	2.1386	1.9901
	(1.894;2.1178)	(2.0409;2.2291)	(1.8671;2.0991)
α_{01}	35.7973	40.2038	34.9343
	(31.828;39.522)	(36.529;43.839)	(30.9276;38.8261)
α_{11}	1.1591	1.2795	1.261
	(0.1957;2.0586)	(0.0636;2.3639)	(0.3938;2.1195)
β_2	-0.1911	0.0152	-0.0706
	(-0.596;0.1749)	(-0.344;0.3395)	(-0.4739;0.2840)
α_{02}	5.9897	6.6010	6.2707
	(4.4525;7.8504)	(4.955;8.5088)	(4.6507;8.198)
α_{12}	3.7613	4.2323	4.0365
	(1.9208;6.4235)	(2.3361;6.7781)	(2.1769;6.6339)

Table 6: Bayes Factors B_{lk}.

$B_{21} = -8.064$	$B_{42} = 40.478$
$B_{31} = 31.6502$	$B_{52} = 40.4632$
$B_{41} = 32.414$	$B_{43} = 0.7638$
$B_{51} = 32.3992$	$B_{53} = 0.7490$
$B_{32} = 39.7142$	$B_{54} = -0.0148$

5 Concluding Remarks

The polyhazards models can be effectivelly used for fitting lifetime data in the presence of a vector of covariates. The use of Markov Chain Monte Carlo methods for a Bayesian analysis of this model is a suitable way to get the posterior summaries of interest.

It is important to point out that we consider $k = 2$ hazard components to be fitted by the censored data sets in Section 4. The value $k = 2$ was used after preliminary analysis of Kaplan-Meier nonparametric estimators for the survival function.

Alternatively, we could assume k unknown and to choose k that maximizes the posterior distributions

$$\pi(k|t) = \frac{m_k v_k}{\sum_j m_j v_j}, \tag{21}$$

where

$$m_k = \int f(t|\Psi_k)\, \pi_k(\Psi_k)\, d\Psi_k, \tag{22}$$

Ψ_k represent all parameters in the polyhazard models, π_k is the prior for Ψ_k and v_j is the prior probability that the number of components is j.

A simplification in the evaluation of m_j is given by the BIC (Bayesian Information Criterion) approximation, $\hat{m}_j = n^{d_j/2} L\left(\hat{\Psi}_j\right)$, where d_j is the dimension of Ψ_j and $L\left(\hat{\Psi}_j\right)$ is the likelihood function evaluated at its maximum (Kass and Wasserman, 1995).

References

Basu, S.; Basu, A.P.; Mukhopadhyay, C. (1999). Bayesian Analysis for Masked System Failure Data Using non-identical Weibull Models. Journal of Statistical Planning and Inference, **78**, 255-275.

Berger, J.O.; Sun, D. (1993). Bayesian analysis for the Poly-Weibull Distribution. Journal of the American Statistical Association, **88**, 1412–1418.

Berger, J.O.; Sun, D. (1996). Bayesian Inference for a Class of Poly-Weibull Distributions. In *Bayesian Analysis in Statistics and Econometrics — Essays in Honor of Arnold Zellner*, (Eds. Berry, D. A., Chaloner, K. M. and Geweke, J. K.), pp. 101–113. New York: Wiley.

Best, N. G.; Cowles, M. K.; Vines, S. K. (1996). *CODA: Convergence Diagnosis and Output Analysis Software for Gibbs Sampling Output, Version 0.3.* MRC Biostatistics Unit, Cambridge.

Davison, A. C.; Louzada-Neto, F. (2000). Inference for the Poly-Weibull Model. The Statistician, **49**, 189-196.

Dempster, A.; Laird, N.; Rubin, D. (1977). Maximum Likelihood from Incomplete Data Via the EM Algorithm. Journal of the Royal Statistical Society, B, **39**, 1–38.

Chib, S.; Greenberger, E. (1995). Understanding the Metropolis-Hastings Algorithm. The American Statistician, 49, 327-335.

Cox, D. R. (1972). Regression Models and Life Tables (with discussion). Journal of the Royal Statistical Society, B, **24**, 406-423.

Doornik, J. A. (1999). *Object-Oriented Matrix Programming Using Ox, 3^{rd}* ed. London: Timberlake Consultants Press and Oxford: www.nuff.ox.ac.uk/Users/Doornik.

Gelfand, A. E. (1995). Model Determination Using Sampling-Based Methods. In *Markov Chain Monte Carlo in Practice* (eds. Gilks, W. R.; Richardson, S.; Spiegelhalter, D. J.), pp. 145-161. London: Chapman and Hall.

Gilks, W. R. (1992). Derivative-Free Adaptive Rejection Sampling for Gibbs Sampling. In *Bayesian statistics 4* (eds. Bernardo, J. M.; Berger, J. O.; Dawid, A. P.; Smith, A. F. M.), pp. 641-649. Oxford: Clarendon Press.

Gilks, W. R. (1996). Full conditional distributions. In *Markov Chain Monte Carlo in Practice* (eds. Gilks, W. R.; Richardson, S.; Spiegelhalter, D. J.), pp. 75-88. London: Chapman and Hall.

Gilks, W. R.; Best, N. G.; Tan, K. K. C. (1995). Adaptive Rejection Metropolis Sampling within Gibbs Sampling. Applied Statistics, **44**, 455-473.

Gilks, W. R.; Neal, R. M.; Best, N. G.; Tan, K. K. C; (1997). Adaptive Rejection Metropolis Sampling (vol 44, pg 455, 1995) (correction or addition). Applied Statistics - Journal of the Royal Statistical Society, C, **46**, 541-542.

Gilks, W. R.; Wild, P. (1992). Adaptive Rejection sampling for Gibbs Sampling. Applied Statistics - Journal of the Royal Statistical Society, C, **41**, 337-348.

Jeffreys, H. (1961). *Theory of Probability, 3^{rd}* ed. Oxford University Press.

Kass, R. E.; Wasserman, L. (1995). A Reference Bayesian Test for Nested Hypothesis and its Relationship to the Schwarz Criterion. Journal of the American Statistical Association, **90**, 928-934.

Klein, J.P.; Moeschberger, M.L. (1997). *Survival Analysis: Techniques for Censored and Truncated Data.* New York: Springer-Verlag.

Kuo, L.; Yang, T. Y. (2000). Bayesian Reliability Modeling for Masked System Lifetime Data. Statistics and Probability Letter, **47**, 229-241.

Lagakos, S. W.; Louis, T. A. (1988). Use the Tumour Lethality to Interpret Tumorigenicity Experiments Lacking Cause-of-Death Data. Applied Statistics, **37**, 169-179.

Louzada-Neto, F. (1999). Polyhazard Regression Models for Lifetime Data. Biometrics, **55**, 1281–1285.

Tanner, M.; Wong, W. (1987). The Calculations of Posterior Distributions by Data Augmentation. Journal of the American Statistical Association, **82**, 528-550.