Inference on the Univariate Frailty Model: An Approach Bayesian Reference Analysis

Vera Lucia D. Tomazella and Camila Bertini Martins
Universidade Federal de São Carlos, SP-Brasil
(vera@ufscar.br; cacamila-bertini@yahoo.com.br)
Jose Miguel Bernardo
Universitat de Valencia, Spain
(jose.m.bernardo@uv.es)

- ABSTRACT: In this work we present an approach involving objective Bayesian reference analysis to the Frailty model with survival time univariate and sources of heterogeneity that are not captured by covariates. The derivation unconditional hazard and survival leads to the result known as Lomax distribution, also known as the Pareto distribution of the second kind. This distribution has an important position in the test of life to adjust data from failures of business. Reference analysis, introduced by Bernardo (1979) produce a new solution for this problem. The results are illustrated with survival data analyzed in the literature and simulated data.

- **Key Words**: Reference analysis, Univariate data, Frailty model, Unobserved heterogeneity

1 Introduction

The notion of frailty provides a convenient way to introduce random effects, dependence and unobserved heterogeneity into models for survival data. In its simplest form, a frailty is an unobserved random proportionality factor that modifies the hazard function of an individual, or of related individuals. One can distinguish two broad classes of frailty models: models with an univariate survival time as endpoint and models that describe multivariate survival endpoints (e.g., competing risks, recurrence of events in the same individual, occurrence of a disease in relatives). The term frailty itself was introduced by Vaupel et al. (1979) in univariate survival models and the model was substantially promoted by its application to multivariate survival data in a seminal paper by Clayton (1978) (without using the notion "frailty") on chronic disease incidence in families. Frailty models are extensions of the proportional hazards model which is best known as the Cox model (Cox, 1972), the most popular model in survival analysis. Normally, in most clinical applications, survival analysis implicitly assumes a homogeneous population to be studied. This means that all individuals sampled into that study are subject in principle under the same risk (e.g., risk of death, risk of disease recurrence). In many applications, the study population can not be assumed to be homogeneous but must be considered as a heterogeneous sample, i.e. a mixture of individuals with different hazards. For example, in many cases it is impossible to measure all relevant covariates related to the disease of interest, sometimes because of economical reasons, sometimes the importance of some covariates is still unknown.

The frailty approach is a statistical modelling concept which aims to account for heterogeneity, caused by unmeasured covariates. In statistical terms, a frailty model is a random effect model for time-to-event data, where the random effect (the frailty) has a multiplicative effect on the baseline hazard function.

Univariate lifetime is used to describe the influence of unobserved covariates in a proportional hazards model (heterogeneity). The variability of survival data is split into a part that depends on risk factors, and is therefore theoretically predictable, and a part that is initially unpredictable, even when all relevant

information is known. A separation of these two sources of variability has the advantage that heterogeneity can explain some unexpected results or give an alternative interpretation of some results, for example, Aalen and Tretli (1999) analyzed the incidence of testis cancer by means of a frailty model based on data from the Norwegian Cancer Registry collected during 1953-93.

One important problem in the area of frailty models is the choice of the frailty distribution. The frailty distributions most often applied is the gamma distribution (Clayton 1978; Vaupel et al. 1979). The model presented in this article will be studied with the gamma frailty distribution under a objective Bayesian perspective. Sinha and Dey (1997) and Tomazella et all. (2004) are some Bayesian work with frailty model.

From a Bayesian perspective, the outcome of any inference problem is the posterior distribution of the quantity of interest, which combines the information provided by the data with available prior information; it has been often recognized that there is a pragmatically important need for a form of prior to posterior analysis which captures, in a well-defined sense, the notion that the prior should have a minimal effect, relative to the data, on the posterior inference.

Reference analysis, introduced by Bernardo (1979) and further developed by Berger and Bernardo (1989, 1992a,b,c) seems to be the only available method to achieve posterior distributions which produces objective Bayesian inference, meaning that inferential statements depend only on the postulated model and the available data. Moreover, there is the requirement that the prior distribution is minimally informative in a precise information-theoretic sense. The information provided by the data should dominate the prior information reflecting the vague nature of the prior knowledge. The driving idea is to maximize the expected Kullback-Leibler divergence of the posterior distribution with respect to the prior. Starting from a reference prior, the reference posterior is a consequence of a formal application of the Bayes theorem. Reference analysis provides posterior distributions with some nice properties, such as generality and invariance, consistent marginalization and consistent sampling properties.

Section 2 we present a summary of reference analysis. In Section 3, the theory is applied to an inference problem, the frailty model parameters. In Section 4 a example simulated data is presented. In Section 5 a real numeric example is presented. In Section 6 include discussion.

2 Reference Analysis

Reference analysis, introduced by Bernardo (1979) and further developed by Berger and Bernardo (1989, 1992a,b,c) appears to be the only available method to derive objective posterior distributions which produces objective Bayesian inference, in the sense that inferential statements depend only on the assumed model and the available data, and the prior distribution used to make an inference is least informative in a certain information-theoretic sense. The reference Bayesian analysis is to specify a prior distribution such that, even for moderate sample sizes, the information provided by the data should dominate the prior information because of the "vague" nature of the prior knowledge. Reference analysis uses the concept of statistical information, in the technical sense of Shannon (1948) and Lindley (1956). To make this notion precise, see Bernardo (2005) for a recent discussion of these ideas.

An important feature in the Berger and Bernardo approach to construct a non-informative prior is the different treatment for interest and nuisance parameters. When there are nuisance parameters (typical case in this paper), one must establish an ordered parameterization with the parameter of interest singled out and then follow the procedure below.

Consider $p(\mathbf{x}|\phi,\lambda)$, $(\phi,\lambda) \in \Phi \times \Lambda \subseteq \Re \times \Re$ a probability model with two real-valued parameters ϕ and λ, where ϕ is the quantity of interest , and suppose that the joint posterior distribution of (ϕ,λ) is asymptotically normal with covariance matrix $S(\hat{\phi},\hat{\lambda})$. So if $S(\phi,\lambda) = H^{-1}(\phi,\lambda)$ we have,

(i) the conditional reference prior of λ given λ is

$$\pi(\lambda|\phi) \propto h_{22}(\phi,\lambda)^{1/2} \qquad \lambda \in \Lambda(\phi)$$

(ii) if $\pi(\lambda|\phi)$ is not proper, a compact approximation $\{\Lambda_i(\phi),\ i = 1, 2...\}$ to $\Lambda(\phi)$ is required, and the reference prior of λ given ϕ is given by

$$\pi_i(\lambda|\phi) = \frac{h_{22}(\phi,\lambda)^{1/2}}{\int_{\Lambda_i(\phi)} h_{22}(\phi,\lambda)^{1/2} d\lambda}, \qquad \lambda \in \Lambda_i(\phi)$$

(iii) within each $\Lambda_i(\phi)$ the marginal reference prior of ϕ is obtained as

$$\pi_i(\phi) \propto \exp\left\{ \int_{\Lambda_i(\phi)} \pi_i(\lambda|\phi) \log\left[s_{11}^{1/2}(\phi,\lambda) \right] d\lambda \right\}$$

where $s_{11}^{1/2}(\phi,\lambda) = h_\phi(\phi,\lambda) = h_{11} - h_{12}h_{22}^{-1}h_{21}$

(iv) the reference posterior distribution of ϕ given data $\{x_1, ..., x_n\}$ is

$$\pi(\phi|x_1,...,x_n) \propto \pi(\phi) \left\{ \int_{\Lambda(\phi)} \left\{ \prod_{l=1}^{n} p(x_l|\phi,\lambda) \right\} \pi(\lambda|\phi)\, d\lambda \right\}$$

Under appropriate regularity conditions (see Bernardo (2005) for more details) the joint reference prior can be written as the product of two independent functions of parameters as follows.

If the nuisance parameter space $\Lambda(\phi) = \Lambda$ is independent of ϕ, and the functions $s_{11}^{-1/2}(\phi,\lambda)$ and $h_{22}^{1/2}(\phi,\lambda)$ factorize in the form

$$\{s_{11}(\phi,\lambda)\}^{-1/2} = f_1(\phi)\, g_1(\lambda), \qquad \{h_{22}(\phi,\lambda)\}^{1/2} = f_2(\phi)\, g_2(\lambda),$$

then,

$$\pi(\phi) \propto f_1(\phi), \ \ and \qquad \pi(\lambda|\phi) \propto g_2(\lambda).$$

the joint reference prior relative the parametric value ordered (ϕ,λ) is given by

$$\pi(\lambda,\phi) = f_1(\phi)\, g_2(\lambda).$$

In this case, there is no need for compact approximation, even if the conditional reference prior is not proper (see Bernardo and Smith, 1994).

3 The Frailty Model

The standard situation of the application of survival methods in clinical research assumes that a homogeneous population is investigated when subject are under different conditions (e.g., experimental treatment and standard treatment). The appropriate survival model then assumes that the survival data of the different patients are independent form each other and that each patient's individual survival time distribution is the same (independent and identically distributed failure times). This basic presumption implies a homogeneous population. However, in the field of clinical trials, one observes in many most practical situations that patients differ substantially. The effect of a drug, a treatment, or the influence of various explanatory variables may differ greatly between subgroups of patients.

Univariate frailty models take into account that the population is not homogeneous. Heterogeneity may be explained by covariates, but when important covariates have not been observed, this leads to unobserved heterogeneity. Vaupel et al. (1979) introduced univariate frailty models (with a gamma distribution) into survival analysis to account for unobserved heterogeneity or missing covariates in the study population. The idea is to suppose that different patients possess different frailties and patients more "frail" or "prone" tend to have the event earlier that those who are less frail.

In this work we consider unobserved source of heterogeneity that are not readily captured by covariate. The univariate frailty model extents the Cox model such that the hazard of an individual depends in addition

on an unobservable random variable v, which acts multiplicatively on the baseline hazard function, so the hazard for individual i at time t is

$$h_i(t) = h_i(t|\nu_i) = \nu_i h_0(t), \tag{1}$$

where ν_i represent the individuals frailty, and a baseline hazard $h_0(t)$. The individual hazard $h_i(t)$ is interpreted as a conditional hazard given ν_i. The model (26) is sometimes refereed as Clayton's multiplicative model (Clayton,1991), where the frailty term represents the information that may not be observed such as, environment and genetics factors or information that, by some reason, were not considered at the planning. An important point is that the frailty ν is an unobservable random variable varying over the sample which increases the individual risk if $\nu > 1$ or decreases if $\nu > 1$.

The corresponding conditional survival function is

$$S_i(t) = S_i(t|\nu_i) = S_0(t)^{\nu_i}, \tag{2}$$

representing the probability of being alive at t given the random effect ν_i.

Note that until now the model is described at individual level, but this individual model is not observable. That is the reason why it is essential to consider the model at a population level. The survival of the total population is the mean of the individual survival functions.

An important point is the identifiability of univariate frailty models. Univariate frailty models are not identifiable from the survival information alone. However, Elbers and Ridder (1982) proved that a frailty model with finite mean is identifiable with univariate data, when covariates are included in the model.

The frailty term is random and we must assume a distribution for it. Natural frailty distribution candidates supposed to be continue and time dependent, are the gamma, log-normal, inverse Gaussian an Weibull. In this work we consider the gamma. This distribution has the range $(0, \infty)$ and is mathematically tractable. In order to keep the identifiability of the model it is convenient to take the distribution with mean 1 so we assume a gamma distribution with parameter $(\frac{1}{\alpha}, \frac{1}{\alpha})$, where α quantifies the amount of heterogeneity among subjects.

3.1 Unconditional Hazard and Survival

To obtain the unconditional survival function we need to integrate out the frailty term.

$$S(t) = \int_0^\infty S(t|\nu) g(\nu) d\nu. \tag{3}$$

To obtain the unconditional hazard we start from the unconditional survival and take negative logs to obtain the cumulative hazard

$$\begin{aligned} H(t) &= -\log S(t) && (4)\\ &= -\log \int_0^\infty S(t|\nu) g(\nu) d\nu. && (5) \end{aligned}$$

The very useful tool in frailty analysis is the *Laplace Transform*. Given a function $f(x)$, the Laplace Transform considered as function of a real argument s is defined as

$$Q(s) = \int_0^\infty e^{sx} f(x) dx. \tag{6}$$

The reason why this is useful in our context is that the Laplace transform has exactly the same form as the unconditional survival function. Think of $f(x)$ as the frailty distribution $g(\nu)$ and s as the cumulative baseline hazard $H_0(t)$ and we obtain

$$S(t) = \int_0^\infty e^{H_0(t)\nu} f(\nu) d\nu = Q(H_0(t)), \tag{7}$$

where Q denotes the Laplace transform.

Vaupel (1990) has defined the frailty transform as the function

$$F(m, s) = \int_0^\infty \nu^m e^{s\nu} g(\nu) d\nu \tag{8}$$

Note que $F(0, s)$ is the Laplace transform (6), and $F(m, s)$ give the m-th moment of the distribution of ν at birth.

The Laplace transform of the gamma distribution with parameters $(\frac{1}{\alpha}, \frac{1}{\alpha})$, is

$$Q(s) = \left(\frac{\frac{1}{\alpha}}{\frac{1}{\alpha} + s} \right)^{\frac{1}{\alpha}}. \tag{9}$$

Evaluating this at $s = H_0(t)$ we obtain the result for survival function

$$S(t) = \frac{1}{(1 + \alpha H_0(t))^{\frac{1}{\alpha}}} \tag{10}$$

The unconditional (population) hazard function under gamma frailty is

$$h(t) = \frac{h_0(t)}{1 + \alpha H_0(t)}. \tag{11}$$

For the baseline hazard function we using the exponential distribution. This distribution has been extensively used to model the baseline hazard function due to its simplicity and flexibility. This is the particular case where the hazard is constant for each individual. So if the hazard baseline is $\lambda_0(t) = \lambda$ and frailty has a gamma distribution the population hazard is

$$h(t) = \frac{\lambda}{1 + \alpha\lambda t}, \tag{12}$$

where λ is the average individual hazard and α is the variance of frailty.

The result is know as a Pareto distribution of the second kind or Lomax distribution (Lomax, 1954) with parameters $(\frac{1}{\alpha}, \alpha\lambda)$, where its probability density function (pdf) is

$$f(t) = \frac{\lambda}{(1 + \alpha\lambda t)^{\frac{1}{\alpha}+1}}, \quad t > 0,\ \alpha > 0,\ \lambda > 0. \tag{13}$$

The survival function associated with (13) is given by

$$S(t) = \frac{1}{(1 + \alpha\lambda t)^{\frac{1}{\alpha}}}. \tag{14}$$

Further, $E(t) = \left[\alpha\lambda(\frac{1}{\alpha} - 1)\right]^{-1}$, $\alpha < 1$ and $Var(t) = \frac{1}{\alpha}\left[\alpha\lambda(\frac{1}{\alpha} - 1)(\frac{1}{\alpha} - 2)\right]^{-1}$, $\alpha < \frac{1}{2}$.

The Lomax distribution, occurs in economics (Arnold, 1983) and reliability theory (Harris, 1968). In reliability theory it appears as a mixture of the one parameters exponential.

In the particular case where $\alpha = 1$ the Figure 1 show the density function (13), survival function (14) and hazard function (12) for value different of λ

Let $t_1, \ldots, t_n$ be a random sample of lifetime data, the likelihood function for n individual is given by

$$L(\lambda, \alpha) = \prod_{i=1}^{n} h(t_i) S(t_i) \tag{15}$$

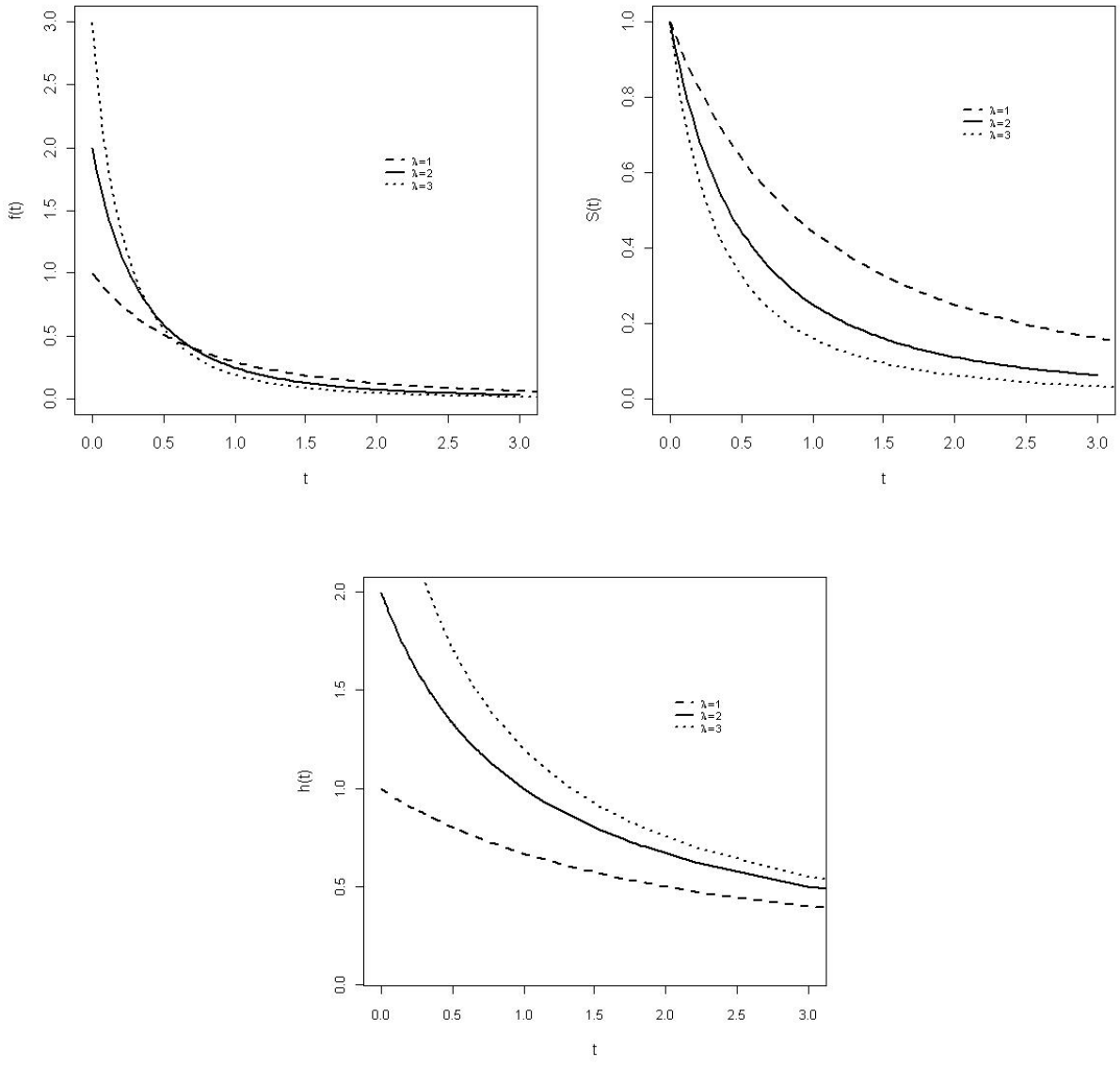

Figure 1: Density function (13),survival function (14) and hazard function (12)

Considering (10) and (12) have

$$L(\lambda,\alpha)=\prod_{i=1}^{n}\frac{\lambda}{(1+\alpha\lambda t_i)^{1/\alpha+1}}. \tag{16}$$

The Figure 2 show the behaviour of the likelihood (16).

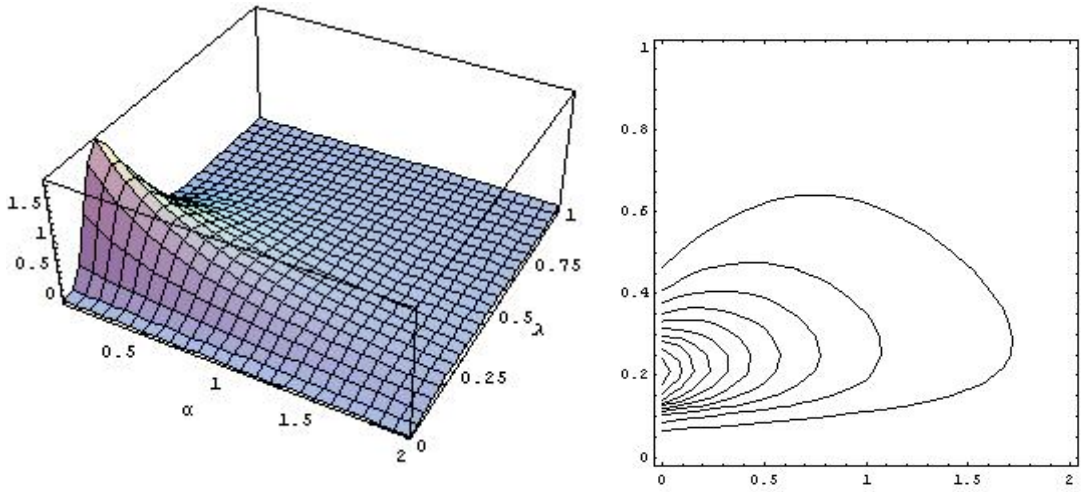

Figure 2: likelihood function (16)

The log-likelihood function(16) is given by

$$l(\lambda,\alpha)=n\log\lambda-(1/\alpha+1)\sum_{i=1}^{n}\log(1+\alpha\lambda t_i). \tag{17}$$

Interval estimates and hypothesis tests for the parameters can be performed, in principle, by considering the asymptotic normal distribution of the maximum likelihood estimates (mles) and the asymptotic chi-squared distribution of the likelihood ratio statistics, respectively (Lawless,1982). As these procedures are standard we turn to the Bayesian approach.

3.2 Reference Analysis for the frailty parameter

The posterior distribution of the parameter is often asymptotically normal (see e.g., Bernardo and Smith (1994, Sec. 5.3). In this case, the reference prior is easily derived. The posterior distribution is asymptotically normal, than reference prior only depends on Fisher information matrix. In this section we derive the reference prior considering the approach of one nuisance parameter described in the section (2). Considering that in this model we have preference by parameter α that quantifies the amount of heterogeneity among subjects, the parameters λ is nuisance parameter.

Considering the log of the likelihood function (17) we have that the first and second derivatives are given by

$\frac{\partial l}{\partial\lambda}=\frac{n}{\lambda}-(1/\alpha+1)\sum_{i=1}^{n}\frac{\alpha t_i}{1+\alpha\lambda t_i}$

$\frac{\partial l}{\partial\alpha}=\frac{1}{\alpha^2}\sum_{i=1}^{n}\log(1+\alpha\lambda t_i)-\frac{1+\alpha}{\alpha}\sum_{i=1}^{n}\frac{\lambda t_i}{1+\alpha\lambda t_i}$

$\frac{\partial^2 l}{\partial \lambda^2} = \frac{n}{\lambda} - (1/\alpha + 1)\sum_{i=1}^{n} \frac{\alpha t_i^2}{(1+\alpha\lambda ti)^2}$

$\frac{\partial^2 l}{\partial \alpha^2} = \frac{2}{\alpha^3}\sum_{i=1}^{n} \log(1+\alpha\lambda t_i) - (1 + \frac{1}{\alpha}\sum_{i=1}^{n} \frac{\lambda^2 t_i}{(1+\alpha\lambda t_i)^2} - \frac{2}{\alpha^2}\sum_{i=1}^{n} \frac{\lambda t_i}{1+\alpha\lambda t_i}$

$\frac{\partial^2 l}{\partial \alpha \partial \lambda} = -\frac{1}{\alpha^2}\sum_{i=1}^{n} \frac{\alpha t_i}{1+\alpha\lambda t_i} - \sum_{i=1}^{n} \frac{\lambda \alpha t_i^2}{(1+\alpha\lambda t_i)^2} + (1/\alpha + 1)\sum_{i=1}^{n} \frac{t_i}{1+\alpha\lambda t_i}$

Setting the results of the first partial derivatives equal to zero, an iteration procedure may be used to obtain the estimate of α and λ. The Fisher's matrix is given by

$$H(\alpha,\lambda) = \begin{bmatrix} \frac{2+\alpha(3+\alpha-4\alpha^3-2\alpha^4+\alpha^6+\alpha^7)}{\alpha^6(\alpha^6(1+\alpha)(2+\alpha))} & -\alpha\left[\frac{\alpha}{\lambda(2+\alpha)} - \frac{1}{\alpha^2\lambda(1+\lambda)}\right] \\ -\alpha\left[\frac{\alpha}{\lambda(2+\alpha)} - \frac{1}{\alpha^2\lambda(1+\lambda)}\right] & \frac{\alpha^3}{\lambda^2(2+\alpha)} \end{bmatrix} \tag{18}$$

The corresponding inverse Fisher's matrix is given by

$$S(\alpha,\lambda) = \begin{bmatrix} \alpha^6(1+\alpha)^2 & -\alpha^2(1+\alpha)(-2-\alpha+\alpha^3+\alpha^4) \\ -\alpha^2(1+\alpha)(-2-\alpha+\alpha^3+\alpha^4) & \frac{\lambda^2(1+\alpha)}{\alpha^3}\left[2+\alpha(3+\alpha-4\alpha^3-2\alpha^4+\alpha^6+\alpha^7)\right] \end{bmatrix} \tag{19}$$

Hence $h_{\lambda\lambda}^{1/2}(\alpha,\lambda)$ and $S_{\alpha\alpha}^{-1/2}(\alpha,\lambda)$ factorize, we have

$$h_{\lambda\lambda}^{1/2}(\alpha,\lambda) \propto \sqrt{\frac{\alpha^3}{2+\alpha}}\frac{1}{\lambda} = f_\lambda(\alpha)g_\lambda(\lambda) \tag{20}$$

and

$$s_{\alpha\alpha}^{-1/2}(\alpha,\lambda) \propto \frac{1}{\alpha^3(1+\alpha)} = f_\alpha(\alpha)g_\alpha(\lambda) \tag{21}$$

So that reference prior of λ given α is

$$\pi(\lambda|\alpha) \propto \frac{1}{\lambda} \tag{22}$$

and the marginal reference prior of the parameter of interest α, is given by

$$\pi(\alpha) \propto \frac{1}{\alpha^3(1+\alpha)}. \tag{23}$$

Combining (22) and (23), the joint reference prior needed to obtain a reference posterior for the parameter of interest α is given by

$$\pi(\alpha,\lambda) \propto \pi(\alpha)\pi(\lambda|\alpha) = \frac{1}{\lambda\alpha^3(1+\alpha)}. \tag{24}$$

The Figure 3 show the reference prior for α (23) and the joint reference prior for α and λ (24).

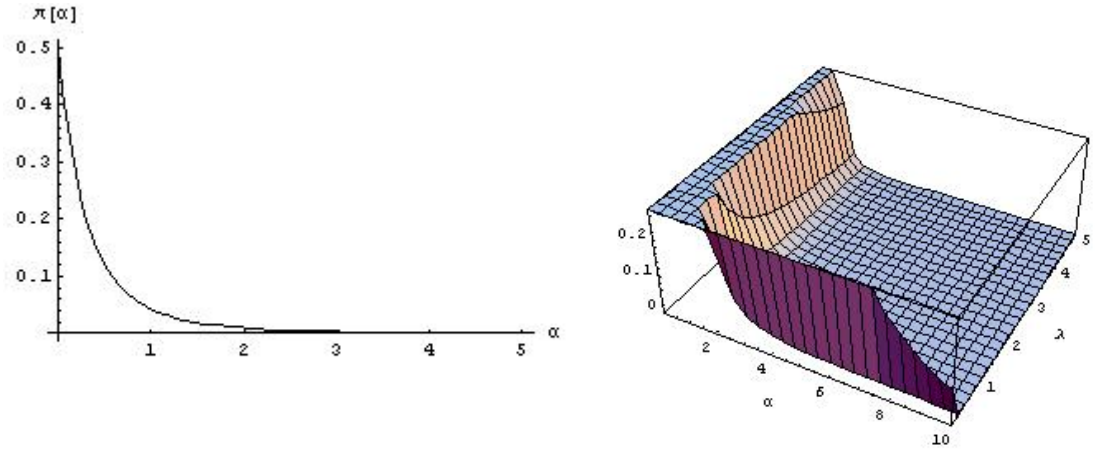

Figure 3: Reference prior for α (23) (left) and joint reference prior (24) (right)

Therefore, using (24) in Bayes theorem to derive the corresponding joint posterior and integrating out the nuisance parameter λ, the reference posterior for the parameter α

$$\pi(\alpha|t_1,\ldots,t_n) \propto \pi(\alpha)\int_0^\infty \lambda^{(n-1)}\prod_{i=1}^{n}\frac{1}{(1+\alpha\lambda t_i)^{(1/\alpha+1)}}d\lambda. \tag{25}$$

Defined the desired reference posterior distribution of the quantity of interest α the next step is to obtain inference of the parameters. In our case, can not be obtained the marginal posterior densities explicitly. We overcome this difficulty by making use of the Markov Chain Monte Carlo (MCMC) methodology to obtain approximations for such densities.

4 Example with simulated data

As a brief numerical illustration of posterior densities given in (25), we consider four simulated samples with $n = 10$, $n = 25$, $n = 50$ and $n = 100$, observation from the Lomax distribution with $\alpha = 0.25$, $\lambda = 4$.

Considering the simulated data set it was obtained a samples of the reference posterior (25) generated by the Metropolis-Hastings technique (Hastings, 1970), i.e., through Markov Chain Monte Carlo methods implemented in software R. The convergence of the chains were assessed according to convergence diagnostics implemented in CODA (Best *et al.*, 1997). Graphical traces of those methods and kernel density estimation for the parameter α and λ showed that there were no convergence problems. We generated two chains of $50,000$ iterations each for the model parameters.

The posterior results are summarized in Table 1, where it is presented characteristics such as posterior mean and standard deviation (SD) and the credible intervals of (95%) of the parameters of interest considering the simulated samples with $n = 10$, $n = 25$, $n = 50$ and $n = 100$.

In the Figures 4, show plots of the generated samples $n = 10$ and $n = 100$, where also we observed equilibrium around of one point. The Figure 5 show the empirical marginal posteriors for model parameter α and λ based on the generated chains of the marginal reference posterior (25).

From the results simulated, we conclude that the prior distributions yield posterior densities adequately. It may be observed that even with rather small and moderate samples sizes, the reference posterior probabilities are not appreciably different (see Figure 5). However a comparative study in terms of frequentist coverage probabilities of this priors is needed.

Table 1: estimated parameters for simulated sample

	α			λ		
n	Mean	SD	Cred. Int.(95%)	Mean	SD	Cred. Int.(95%)
10	$0,254$	$0,0135$	$[0,230\ ;\ 0,282]$	$3,96$	$0,36$	$[3,33\ ;\ 4,71]$
25	$0,255$	$0,0135$	$[0,231\ ;\ 0,282]$	$3,92$	$0,35$	$[3,29\ ;\ 4,65]$
50	$0,254$	$0,0134$	$[0,230\ ;\ 0,282]$	$3,93$	$0,35$	$[3,31\ ;\ 4,66]$
100	$0,256$	$0,0128$	$[0,234\ ;\ 0,283]$	$3,84$	$0,31$	$[3,26\ ;\ 4,48]$

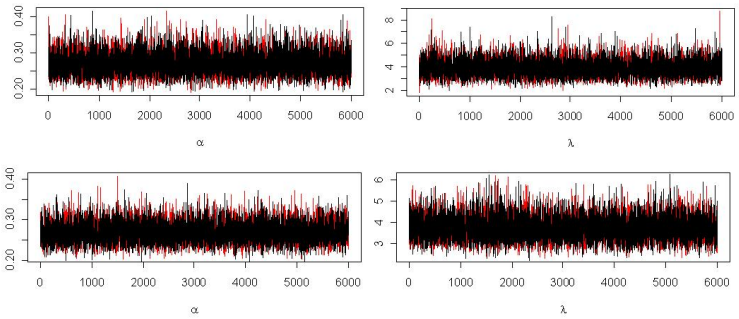

Figure 4: Trace of α and λ for $n = 10$ (above) and for $n = 100$ (down)

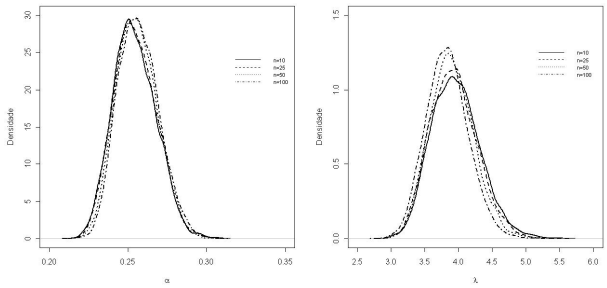

Figure 5: Marginal posterior density of α and λ for all generate sample

5 Example with real data

In this section the methodology is illustrated considering the leukemia survival data introduced by Feigel and Zelen (1965), where the concomitant variable WBC, the white blood cell count of a patient, is not considered in the model. In the Table 2, we have the data of $n = 17$ positive patients and $n = 16$ negative patients. We want to find inference for α frailty variance and λ the average individual hazard.

The posterior results for interest parameters are summarized in Table 3. In the Figures 6, show the empirical marginal posteriors for model parameter α and λ based on the generated chains of the marginal reference posterior.

Figure 7 shows that there is evidence of convergence for both parameters, because the chains they behave within a limit. This can also be seen in Figure 6, since the densities estimated parameters of interest for the two chains are generated close.

It is important to note that there is significant of the frailty variance indicated by the estimated of α which is equal to 2.583 and the average individual hazard is too significant indicated by estimated of λ, which is equal to 0.194.

Table 2: Feigl and Zelen survival time data (Weeks)

Times	Ag+	$\log_{10}$ (wbc)	Times	Ag-	$\log_{10}$ (wbc)
63	2300	3,36	56	4400	3,64
156	750	3,88	65	3000	3,48
134	2600	3,41	7	1500	3,18
16	6000	3,78	16	9000	3,95
108	10000	4,02	22	5300	3,72
121	10000	4,00	3	10000	4,00
4	17000	4,23	4	19000	4,28
39	5400	3,73	2	27000	4,43
143	7000	3,85	3	28000	4,45
56	9400	3,97	8	31000	4,49
26	32000	4,51	4	26000	4,41
22	35000	4,54	3	21000	4,32
1	100000	5,00	30	79000	4,9
1	100000	5,00	4	10000	5,00
5	52000	4,72	43	10000	5,00
65	100000	5,00			

Table 3: Posterior results of (25)

Parameters	Mean	SD	Cred.Int.(95%)
α	2.583	0.423	$[1.89; 3.54]$
λ	0.194	0.044	$[0.12; 0.29]$

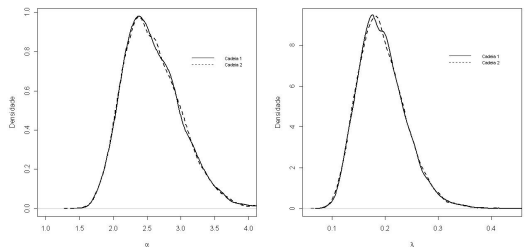

Figure 6: Marginal posterior density α and λ

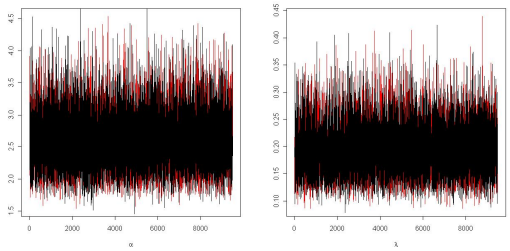

Figure 7: Trace for α and λ

6 Discussion

In this paper we have summarized the derivation of the reference prior distributions and we have illustrated with an important example in survival analysis, where the results of this study, however, suggest that ignoring

frailty may result in biased estimates, considering that the variance of frailty is significant.

Although we considered the gamma frailty distribution and exponential lifetime distribution, this methodology is, in principle, general and can be applied to other frailty distributions, such as inverse Gaussian, log-normal and Weibull distributions, as well as other lifetime distributions. However, mathematical difficulties may be found, even using MCMC Methods to infer over the model parameters.

One possible solution to this problem is to consider a model that captures the effect of covariates. The general model we will entertain is a proportional hazards model with a frailty term where the hazard at time t for an individual with covariates x and frailty ν is,

$$h(t, x, \nu) = \nu h_0(t) \exp(x^{'}\beta), \tag{26}$$

where the distribution of ν does not depend on the observed covariate x.

7 Reference

Aalen, O.O., Tretli, S. (1999) Analysing incidence of testis cancer by means of a frailty model. Camcer Causes and Control 10, 285-292.

Arnold, B. (1983). Pareto Distribution. International Co-operative Publishing House, Burtonsville, M.D.

Berger, J. O. and Bernardo, J. M. (1992a). Ordered group reference priors with applications to a multinomial problem. Biometrika 79, 25–37.

Berger, J. O. and Bernardo, J. M. (1992b). Reference priors in a variance components problem. Bayesian Analysis in Statistics and Econometrics (P. K. Goel and N. S. Iyengar, eds.) Berlin: Springer, 323–340.

Berger, J. O. and Bernardo, J. M. (1992c). On the development of reference priors. Bayesian Statistics 4 (J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith, eds.) Oxford: University Press, 35–60 (with discussion).

Bernardo, J. M. (1979). Reference posterior distributions for Bayesian inference.J. Roy. Statist. Soc. B 41, 113–147 (with discussion). Reprinted in Bayesian Inference (N. G. Polson and G. C. Tiao, eds.) Brookfield, VT: Edward Elgar, 1995, 229–263.

Bernardo, J. M. (1980). A Bayesian analysis of classical hypothesis testing. Bayesian Statistics (J. M. Bernardo, M. H. DeGroot, D. V. Lindley and A. F. M. Smith, eds.) Valencia: University Press, 605–647 (with discussion).

Bernardo, J.M. (2005). Reference analysis, in: D.K. Dey and C.R. Rao, eds., Handbook of Statistics, Vol. 25 (Elsevier, Amsterdam) 17-90.

Bernardo, J. M. and Smith, A. F. M. (1994). Bayesian Theory. Chichester: Wiley. Blyth, S. and Smith, A. F. M. (1998). Bayesian meta-analysis of frequentist p-values. Comm. Statist. Theory and Methods 27, 2707–2724.

Best, N., Cowles, M. and Vines, K. (1997). *CODA - Convergence Diagonosis and Output Analysis software for Gibbs sampling output*: Version *0.4*, MRC Biostatistics Unit,

Chib, S. and Greenberg, E. (1995). Understanding the Metropolis-Hastings Algorithm, The American Statistician, **49**, 327–335.

Clayton, D. G. (1978). A model for association in bivariate life tables and its application in epidemological studies of familial tendency in chronic diease incidence, Biometrika, 65: 141-151.

Cox D.R. (1972a). Regression models and life tables (with discussion), Journal of the Royal Statistical Society B 34: 187-220.

Cox D.R. (1972b).The statistical analysis of dependencies in point processes. In Stochastic point Processes, Ed. P.A.W. Lewis, pp.55-66. New

Elbers, C. and Ridder, G. (1982). True and spurious duration dependence: The identificability of the proportional hazard model. Revew of Economic Studies XLIX, 403-409.

Feigl, P. and Zelen, M. (1965). Estimation of exponential survival probabilities with concomitant Information. Biometrics,21, 826-837.

Harris, O.R. (1968). O. R. 16 pp 307-13.

Hastings, W.K. (1970). Monte Carlo sampling methods using Markov chains and their applications, *Biometrika*, **57**, 97–109.

Lawless, J.F. (1982). Statisctical Models and Methods for Life Time Data. New York: Jonh Wiley.

Lawless, J.F. (1995). The analysis of recurrent events for multiple subjects. Applied Statistics 44, 487-98.

Lindley, D. V. (1956). On a measure of information provided by an experiment. Ann. Math. Statist. 27, 986-1005.

Shannon, C. E. (1948). A mathematical theory of communication. Bell System Tech. J. 27 379-423 and 623-656. Reprinted in The Mathematical Theory of Communication (Shannon, C. E. andWeaver, W., 1949). Urbana, IL.: Univ. Illinois Press.

Sinha, D. and Dey, D. K. (1997). Semiparametric Bayesian analysis of survival data. Journal of the American Statistical Association, 92, 1195-1212.

Tomazella, V.L.D; Louzada, F.N.and Andrade M.G.(2004).Bayesian modelling for multivariate lifetime data with a homogeneous Poisson process with a frailty term. Brazilian Journal of Probability and Statistics, 18, pp. 19-35.

Vaupel, J. W. , Manton, K. G. and Stallard, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography,* 16, 439-454.