Chilean Journal of Statistics
Vol. 4, No. 2, September 2013, 95–109

SURVIVAL ANALYSIS
RESEARCH PAPER

The effect of frailty term in the standard mixture model

Vinicius F. Calsavara[1], Vera L. D. Tomazella[2,*] and José C. Fogo[2]

[1]Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo-SP, Brazil
[2]Departamento de Estatística, Universidade Federal de São Carlos, São Carlos-SP, Brazil

(Received: 08 September 2011 · Accepted in final form: 21 March 2013)

Abstract

In this paper we present the standard mixture model with a frailty term. This model has the advantage in relation the usual survival models, in the sense of incorporate the heterogeneity of two subpopulations (immune and susceptible) to the event of interest. We consider the model with a frailty term as alternative to modelling long-term survival data. In this model a random effect, also called frailty, is introduced on the baseline hazard function, in order to control the unobservable heterogeneity of the units under study. A simulation study is present to analyze the frequentists properties of estimation procedures. A cost study in the estimation of the cured rate is discussed. We explore the use of the `gamlss` package in R as a tool for inference in cure rate models. The developed procedures are illustrated on a real data set.

Keywords: Frailty model · GAMLSS · Long-term survival models · Survival analysis · Variances ratio.

Mathematics Subject Classification: Primary 62N01 · Secondary 62P10.

1. Introduction

In standard survival models, there is an assumption that all units involved in the experiment are at risk and certainly it will experience the event of interest. However, consider the usual survival models to certain data sets can not be appropriate. There are survival data in which a portion of units under study will never experience the event of interest, even if accompanied by a sufficiently large time. Certainly a bulb will fail, however, a patient "cured" of cancer can not experience the disease recurrence. We say that these individuals are (immune, cured or non-susceptible) to the event of interest and the population which they belong has a cure fraction. In this context, the standard mixture model by Boag (1949) and Berkson and Gage (1952) is the best known to estimate the cured rate.

The cure rate models allow a survivors proportion in the population. These models have historically been utilized to analyze long-term survival data and they allow flexibility in estimate the effects of covariates that can influence in the cured fraction, as well as the

*Corresponding author. Email: vera@ufscar.br

ISSN: 0718-7912 (print)/ISSN: 0718-7920 (online)
© Chilean Statistical Society – Sociedad Chilena de Estadística
http://www.soche.cl/chjs

population at risk. When the cure fraction is not present, the data analysis reduces to the usual techniques of survival data.

The long-term survival models have the advantage in relation to usual survival models, in sense of incorporate the heterogeneity of two subpopulations. Among these models, the most common type is the mixture model, where it considers that the population is divided into two subpopulations (immune and susceptible) to the event of interest.

We find several applications of long-term survival models in areas such as biomedical studies, financial, criminology, demography, reliability, among other. For instance, in biomedical data an event of interest can be the patient's death, which can occur due to tumor recurrence. When we consider financial data, an event of interest can be the shutdown of a customer of the bank due to various reasons. In criminology data, the event of interest can be the crime recurrence. These examples can be found in Anscombe (1961), Farewell (1977), Goldman (1984), Broadhurst and Maller (1991), Meeker and Escobar (1998). For more details and examples of long-term survival models can be found in Maller and Zhou (1996).

The cure rate models implicitly assume for the susceptible individuals a homogeneous population, however, a way of quantifying the amount of observed heterogeneity is including covariate in the model. Like that, a portion of the heterogeneity can be explained by covariates, however there is a degree of heterogeneity induced by unobservable risk factors. The models which incorporate unobservable heterogeneity among individuals are known as frailty models; see Vaupel et al. (1979).

The frailty models are characterized by the inclusion of a random effect, that is, an unobservable random variable, that represents the information that can not be observed or have not been observed, such as: environmental factors, genetic or information that for some reason were not considered at the planning. One of the ways to incorporate this random effect, called frailty, is to introduce on the baseline hazard function, in order to control the unobservable heterogeneity of units under study.

Several authors studied the use of multiplicative frailty models, which represent a generalization of the Cox (1972) proportional hazards model. Andersen (1993) and Hougaard (1995) presented a review of the multiplicative frailty models in the classical perspective, while Sinha and Dey (1997) developed a full review of these models, under the Bayesian point of view. The frailty term was initially introduced by Vaupel et al. (1979) in univariate survival models. Clayton (1978) and Oakes (1982) explored the multivariate survival models. Aalen (1988), Hougaard et al. (1994), Longini Jr and Halloran (1996), Price and Manatunga (2001), Peng et al. (2007), Yau and Ng (2001), Yu and Peng (2008), among others, extended the frailty models considering the cured fraction.

The main goal this paper is to realize a simulation study with the aim of analyzing the frequentists properties of the estimation procedures. Investigate in the presence of frailty term what is the cost in the estimation of the cure rate. An application to a real leukemia data set shows the applicability of the model, which the parameter estimates are obtained through `gamlss` package. The reader is referred to Rigby and Stasinopoulos (2005) for a full account of the GAMLSS framework. Stasinopoulos and Rigby (2007) described the implementation of GAMLSS in R software; see R Development Core Team (2012).

This paper is organized as follows. In Section 2 the frailty model is described. The standard mixture model is presented in Section 3. In Section 4 we present the frailty standard mixture model and we discuss the procedure used to estimate the parameters. A simulation study in Section 5. A study of cost in the estimation of the cured rate is described in Section 6. An application to a real leukemia data set is illustrated in Section 7. Availability of the programs constitute the Section 8. Finally, in Section 9 we make close general remarks.

2. Frailty Model

The multiplicative frailty model is an extent of Cox proportional hazards model, where the individual hazard depends of an unobservable random variable V, which acts multiplicatively on the baseline hazard function, so the hazard for individual i at time t is given by

$$h_i(t) = h_i(t|v_i) = v_i h_0(t) \exp(\mathbf{x}_i^\top \boldsymbol{\beta}), \quad i = 1, \ldots, n, \tag{1}$$

where $v_1, \ldots, v_n$ represent the individual frailties, $h_0(t)$ is baseline hazard function, $\boldsymbol{\beta}$ is a vector of coefficients to be estimated and $\mathbf{x_i}$ is the covariates vector. The frailty model, Equation (1), assumes proportional hazards structure conditioned on the random effect.

Note in Equation (1) that the individual hazard increase if $v_i > 1$, decrease for $v_i < 1$ and when $v_i = 1$ the frailty model reduces to the Cox proportional hazards model.

The key idea of this model is that individuals have different frailties, and that the most frail will die earlier than the less frail, hence the name frailty; see Wienke (2011). The frailty term in this model does not only explain the heterogeneity among individuals, it also allows to assess the covariates effect that for some reason were not considered at the planning. For instance, if an important covariate was not included in the model the unobservable heterogeneity will increase affecting the inferences about the parameters related to the covariates. This way, with inclusion of frailty term in the model it will assuage this problem.

The frailty term is a unobservable random variable and we must assume distribution for it. Due to the way as the frailty term acts on the hazard function, natural frailty distribution candidates supposed to be continue and time independent, such as gamma, lognormal, Weibull and inverse Gaussian distributions; see Hougaard (1995). In many papers, the gamma distribution is often applied, because it presents an easy algebraic treatment.

In the context of proportional hazards, according to Elbers and Ridder (1982) is necessary that the random effect distribution has finite mean for the model to be identifiable. This way, in order to keep the identifiability of the model it is convenient to take the distribution with mean 1. In this paper, we assume a gamma distribution with parameters (α, α) where $1/\alpha$ quantifies the amount of heterogeneity among subjects. If α is large it implies low variability among individuals, that is, a homogeneous population, so the frailty variable values will be equal 1, hence the gamma distribution is degenerate at the point 1, and like that we have the standard proportional hazards model for independent data. On the other hand, if α is small it indicates that there is an unobservable heterogeneity among individuals.

Now let us consider the Equation (1) without the presence of the covariates. The hazard function for ith individual is given by

$$h_i(t) = h_i(t|v_i) = v_i h_0(t),$$

with survival function

$$S_i(t) = S_i(t|v_i) = (S_0(t))^{v_i},$$

where $S_0(t)$ is baseline survival function.

In order to obtain the likelihood function is necessary to find the unconditional survival function, so we need to integrate the frailty term, that is

$$S(t) = \int_0^\infty S(t|v) g(v) dv = \int_0^\infty (S_0(t))^v \, g(v) dv = \int_0^\infty e^{-H_0(t)v} g(v) dv = \mathrm{L}_V(H_0(t)),$$

where $g(\cdot)$ is probability density function of the frailty variable, $H_0(\cdot)$ is the cumulative hazard function and $\mathrm{L}_V(\cdot)$ denotes the Laplace transform. Thus the unconditional survival function considering gamma distribution is given by

$$S(t) = \left(1 + \frac{H_0(t)}{\alpha}\right)^{-\alpha}. \tag{2}$$

From Equation (2), if $\alpha \to \infty$ we have the homogeneous model, that is, $S(t) = S_0(t)$.

3. Mixture model to estimate the cured proportion

The standard mixture model by Berkson and Gage (1952) is the best known in survival analysis to modelling long-term survival data. This model consists of a mixture of parametric distributions being an improper survival function considered for the total population and a proper survival function for the susceptible individuals.

Let T be a nonnegative continuous random variable representing the lifetime of an individual in population. The survivor function of the standard mixture model is given by

$$S_{pop}(t) = p_0 + (1 - p_0)S(t), \tag{3}$$

where p_0 represent the cured rate.

The time until the occurrence of any event of interest, in general, can be accommodated by a probability distribution. In literature, numerous distributions have been used to describe survival times, however, the Weibull distribution is widely used to modelling the survival function of those individuals at risk. Note in Equation (3) as the time increases the improper survival function converges to p_0, exactly the cured fraction, indicating that $S_{pop}(t)$ is an improper function.

The corresponding improper density function is written as

$$f_{pop}(t) = -S'_{pop}(t) = (1 - p_0)f(t),$$

where $f(\cdot)$ is proper density function for the individuals group at risk. The improper hazard function is given by

$$h_{pop}(t) = \frac{f_{pop}(t)}{S_{pop}(t)} = \frac{(1 - p_0)f(t)}{p_0 + (1 - p_0)S(t)}. \tag{4}$$

From Equation (4) the proper hazard function can be written as

$$h(t) = \frac{S_{pop}(t)h_{pop}(t)}{(1 - p_0)S(t)} = \left(\frac{S_{pop}(t)}{S_{pop}(t) - p_0}\right) h_{pop}(t). \tag{5}$$

As $(S_{pop}(t)/(S_{pop}(t) - p_0)) > 1$, we obtain $h_{pop}(t) < h(t)$, that is, the improper hazard function is limited by hazard function. Note in Equation (5) that $h(t)$ does not satisfy the proportional hazards property, because $(S_{pop}(t)/(S_{pop}(t) - p_0))$ always it will depend of t; see Rodrigues et al. (2009). From Equation (4) we have

$$\lim_{t\to\infty} h_{pop}(t) = \lim_{t\to\infty} \frac{(1 - p_0)f(t)}{S_{pop}(t)} = \left(\frac{1 - p_0}{p_0}\right) \lim_{t\to\infty} f(t) = 0. \tag{6}$$

The result in Equation (6) shows us as time increases the population at risk converges to zero, this reflects the fact that the population survival curve stabilize at a value (cured fraction), indicating that a portion of individuals have not failed and possibly they were cured.

4. Frailty standard mixture model

In medical area the argumentation that individuals are different is widely accepted. There are biological variations that are not mensurable among subjects that explaining this fact. This variability is generally considered as one of the most important sources of unobservable variation.

In order to modelling long-term survival data and to quantify the amount of heterogeneity among subjects, the population survival function is given by

$$S_{pop}(t) = p_0 + (1 - p_0)\mathrm{L}_V(H_0(t)), \tag{7}$$

where the unobservable random variable V follows a continuous distribution with support on positive real.

Upon substituting Equation (2) into Equation (7) we obtain the improper survival function

$$S_{pop}(t) = p_0 + (1 - p_0)\left(1 + \frac{H_0(t)}{\alpha}\right)^{-\alpha}, \tag{8}$$

with improper density function

$$f_{pop}(t) = (1 - p_0)h_0(t)\left(1 + \frac{H_0(t)}{\alpha}\right)^{-\alpha-1}.$$

The Equation (8) has the following properties:

- If $p_0 = 0$ and $\alpha \to \infty$, we have homogeneous survival model;
- For $p_0 = 0$, we have gama frailty model;
- When $\alpha \to \infty$, we have standard mixture model.

The frailty standard mixture model comprises some well-known sub-models. This is important because we can perform hypothesis tests to verify whether there are evidence of cured individuals, as well as the unobservable heterogeneity.

4.1 Maximum likelihood estimation

In order to estimate the model parameters, we utilize the method of maximum likelihood. Let us the consider the situation when the time to event is not completely observed and is subject to right censuring. Let C_i denote the censoring time. We observe $\mathrm{T}_i = \min(\mathrm{Y}_i, \mathrm{C}_i)$ and $\delta_i = \mathbb{I}(\mathrm{Y}_i \leq \mathrm{C}_i)$, where $\delta_i = 1$ se Y_i is a time to event and $\delta_i = 0$ if right censored, $i = 1, \ldots, n$. Let $\boldsymbol{\theta}$ denote the parameter vector of the lifetime distribution. For the n pairs of times and censoring indicators $(t_1, \delta_1), \ldots, (t_n, \delta_n)$, the likelihood function under uninformative censoring is given by

$$\mathcal{L}(p_0, \alpha, \boldsymbol{\theta}; \boldsymbol{t}, \boldsymbol{\delta}) \propto \prod_{i=1}^{n} \left(f_{pop}(t_i; p_0, \alpha, \boldsymbol{\theta})\right)^{\delta_i} \left(S_{pop}(t_i; p_0, \alpha, \boldsymbol{\theta})\right)^{1-\delta_i}. \tag{9}$$

Henceforward, we assume a Weibull distribution for the time to event T. In our notation,

$$\mathrm{F}(t;\boldsymbol{\theta}) = 1 - \exp(-\sigma t^{\gamma}) \quad \text{and} \quad f(t;\boldsymbol{\theta}) = \sigma\gamma t^{\gamma-1}\exp(-\sigma t^{\gamma}),$$

where $\boldsymbol{\theta} = (\sigma, \gamma)^{\top}$, with $\sigma > 0$ and $\gamma \geq 0$ are scale and shape parameters, respectively.

Let $\boldsymbol{\vartheta} = (p_0, \sigma, \gamma, \alpha)^{\top}$ the vector of parameters. The estimation of vector $\boldsymbol{\vartheta}$ is done through the maximum likelihood method applied in Equation (9). The numerical maximization of log-likelihood function $\ell(\boldsymbol{\vartheta}) = \log \mathcal{L}(\boldsymbol{\vartheta}; \boldsymbol{t}, \boldsymbol{\delta})$ is realized using by RS method available in the `gamlss` package in R; see Rigby and Stasinopoulos (2005).

The asymptotic properties of maximum likelihood estimators are needed to build confidence intervals and hypotheses tests about the model parameters, utilizing that $\widehat{\boldsymbol{\vartheta}}$ has asymptotic multivariate normal distribution under certain regularity conditions, with mean $\boldsymbol{\vartheta}$ and variance and covariance matrix $\Sigma(\widehat{\boldsymbol{\vartheta}})$, which is estimated by

$$\widehat{\Sigma}(\widehat{\boldsymbol{\vartheta}}) = \left(-\frac{\partial^2 \ell(\boldsymbol{\vartheta}; \boldsymbol{t}, \boldsymbol{\delta})}{\partial\boldsymbol{\vartheta}\partial\boldsymbol{\vartheta}'} \right)^{-1}$$

evaluated at $\boldsymbol{\vartheta} = \widehat{\boldsymbol{\vartheta}}$. The required second derivatives are computed numerically.

5. Simulation study

In view of possibility of a sample there is an amount of unobservable heterogeneity and under premise of homogeneous risk for susceptible individuals, arise the question of the estimators. How do they behave when we ignore the presence of the heterogeneity in the sample?

To answer this question, samples were generated of the frailty standard mixture model, which we consider Weibull distribution for the lifetime. The censoring times were generated independently from the survival times according to an exponential distribution. Initially were generated 1000 samples of the frailty Weibull standard mixture model (FWM), given in Equation (8), following the three scenarios:

- $p_0 = 0.5$, $\sigma = 1$, $\gamma = 1$ and $\alpha = 0.5$;
- $p_0 = 0.5$, $\sigma = 1$, $\gamma = 1$ and $\alpha = 1.0$;
- $p_0 = 0.5$, $\sigma = 1$, $\gamma = 1$ and $\alpha = 2.0$.

For each sample, we fitted the FWM model and Weibull standard mixture model (WM). We considered for initial values in the optimization process the fixed values. The simulations that did not converge were discarded. For both models, we compared mean of the parameter estimates, standard error (SE) and mean square error (MSE). We considered sample sizes $n = 50, 70, 100$ and 300 to assess as it affects the asymptotic properties of the estimators. The simulations results are summarized in Table 1.

The results show that for both models the mean estimates p_0 were not much affected with the increase in sample and in degree heterogeneity. Even for small samples the mean of the estimates were close to the fixed values. An interesting fact occurred on the parameter of cured, the mean of the estimates, mean square error and standard error showed results very similar. This fact is mentioned by Shao and Zhou (2004), which they observed that there was a large difference in deviance, but that estimates of the cured rate did not differ, arguing that the estimate of p_0 is determined largely by data values near the tail to the right, while the likelihood function is influenced evenly throughout the data set.

In the three scenarios studied, as we increase the sample size the mean of the estimates approximate of the fixed values. The inclusion of the frailty term in the standard mixture

Table 1. Simulations results for $p_0 = 0.5$, $\sigma = 1$, $\gamma = 1$ and $\alpha = 0.5, 1$ and 2.

			Degree of heterogeneity $1/\alpha$								
			$\alpha = 0.5$			$\alpha = 1$			$\alpha = 2$		
n	Model	Parameter	Mean	SE	MSE	Mean	SE	MSE	Mean	SE	MSE
50	FWM	p_0	0.510	0.078	0.006	0.507	0.074	0.005	0.507	0.072	0.005
		σ	1.397	1.261	0.910	1.303	0.909	0.638	1.186	0.653	0.328
		γ	1.190	0.460	0.217	1.178	0.389	0.165	1.159	0.349	0.123
		α	0.656	0.601	0.360	1.237	1.207	0.696	2.120	2.444	0.858
	WM	p_0	0.497	0.070	0.005	0.501	0.071	0.005	0.504	0.071	0.005
		σ	0.494	0.132	0.267	0.626	0.155	0.161	0.768	0.177	0.088
		γ	0.446	0.066	0.317	0.660	0.100	0.133	0.837	0.131	0.055
70	FWM	p_0	0.494	0.063	0.004	0.503	0.062	0.004	0.505	0.061	0.004
		σ	1.321	1.016	0.717	1.241	0.748	0.500	1.190	0.560	0.304
		γ	1.138	0.377	0.143	1.132	0.325	0.103	1.131	0.288	0.086
		α	0.647	0.495	0.287	1.289	1.127	0.737	2.120	2.091	0.971
	WM	p_0	0.485	0.059	0.004	0.498	0.060	0.004	0.503	0.060	0.004
		σ	0.495	0.113	0.263	0.628	0.131	0.153	0.766	0.145	0.079
		γ	0.441	0.055	0.318	0.652	0.083	0.133	0.808	0.105	0.053
100	FWM	p_0	0.497	0.052	0.002	0.501	0.052	0.003	0.501	0.051	0.002
		σ	1.239	0.795	0.497	1.190	0.584	0.311	1.109	0.408	0.145
		γ	1.097	0.307	0.093	1.115	0.267	0.077	1.096	0.230	0.050
		α	0.628	0.397	0.195	1.234	0.897	0.605	2.150	1.811	0.831
	WM	p_0	0.490	0.050	0.002	0.496	0.050	0.003	0.500	0.050	0.002
		σ	0.497	0.094	0.258	0.628	0.110	0.147	0.757	0.124	0.075
		γ	0.438	0.045	0.320	0.636	0.067	0.142	0.799	0.087	0.052
300	FWM	p_0	0.502	0.030	0.001	0.501	0.030	0.001	0.501	0.029	0.001
		σ	1.148	0.394	0.151	1.123	0.291	0.099	1.055	0.204	0.040
		γ	1.062	0.169	0.030	1.070	0.148	0.024	1.043	0.125	0.014
		α	0.495	0.158	0.024	0.980	0.375	0.117	1.980	0.961	0.270
	WM	p_0	0.493	0.029	0.001	0.497	0.029	0.001	0.500	0.029	0.001
		σ	0.498	0.054	0.253	0.624	0.063	0.144	0.751	0.071	0.067
		γ	0.405	0.023	0.354	0.596	0.035	0.165	0.761	0.047	0.061

model inflated the SE of the estimates. The estimation of the shape and scale parameters of the Weibull distribution in the model, representing by Equation (3), show lower mean estimates than FWM model. As we increase α (lower heterogeneity), the mean of the estimates σ and γ are close to the fixed value 1, this is expected, since large values of α leads to standard mixture model. As expected, the values of MSE and SE decrease as the sample size increases.

We also obtain for each parameter the asymptotic 95% confidence interval and we verify if the confidence interval contained the true value of the parameter, in order to obtain the coverage probabilities (CP) of the confidence intervals. We also found the mean amplitude. The Table 2 displays the coverage probabilities and mean amplitudes.

The procedure for constructing confidence intervals based on asymptotic normality of maximum likelihood estimators was satisfactory in the most cases. For the parameters p_0 and γ, the simulation results show that the coverage probabilities of confidence intervals based on maximum likelihood estimator are close to the desired confidence level. The same occurs for the parameter σ when we fix $\alpha = 2$, while for $\alpha = 0.5, 1$ and $n \leq 100$ the coverage probabilities are considerably lower than the desired confidence levels. When the sample is small and $\alpha = 0.5, 1$ the coverage probabilities of confidence intervals are considerably lower than the desired confidence levels for parameters σ and α.

Table 2. 95% coverage probabilities for the confidence intervals based on maximum likelihood estimator and mean amplitudes.

		Degree of heterogeneity $1/\alpha$					
		$\alpha = 0.5$		$\alpha = 1$		$\alpha = 2$	
n	Parameter	CP	Amplitude	CP	Amplitude	CP	Amplitude
50	p_0	0.947	0.304	0.953	0.289	0.953	0.283
	σ	0.897	4.941	0.913	3.565	0.959	2.561
	γ	0.959	1.801	0.984	1.523	0.982	1.370
	α	0.888	2.357	0.880	4.733	0.893	9.580
70	p_0	0.935	0.246	0.929	0.243	0.929	0.238
	σ	0.887	3.983	0.911	2.934	0.956	2.195
	γ	0.956	1.478	0.987	1.275	0.991	1.131
	α	0.913	1.942	0.889	4.417	0.876	8.195
100	p_0	0.954	0.203	0.944	0.202	0.946	0.200
	σ	0.898	3.115	0.930	2.289	0.959	1.599
	γ	0.949	1.201	0.980	1.046	0.989	0.902
	α	0.938	1.557	0.885	3.516	0.909	7.097
300	p_0	0.948	0.118	0.941	0.116	0.937	0.115
	σ	0.953	1.543	0.963	1.141	0.980	0.798
	γ	0.967	0.663	0.973	0.581	0.972	0.490
	α	0.938	0.621	0.935	1.471	0.939	3.768

5.1 Artificial data

We consider a sample with 100 observation generated by FMW model following sets of parameters $p_0 = 0.5$, $\sigma = 1$, $\gamma = 1$ and $\alpha = 1$. For the censoring times we suppose an exponential distribution. The generation of the data set and the maximum likelihood estimates were obtained as done in the simulation study. Table 3 displays the parameter estimates for the FWM model which we obtain cure rate of approximately 50%. We also observed that the estimated values are close to the fixed values, as well as the asymptotic 95% confidence intervals contain the true values of the parameters.

Table 3. Parameter estimates for the FWM model following sets of parameters $p_0 = 0.5$, $\sigma = 1$, $\gamma = 1$ and $\alpha = 1$.

Parameter	Estimate	95% confidence interval
p_0	0.502	(0.402 ; 0.602)
σ	1.167	(0.150 ; 2.185)
γ	1.034	(0.508 ; 1.560)
α	1.069	(0.252 ; 4.531)

6. Estimation cost p_0

When we consider frailty models arises the question about the estimation cost of certain parameters with the inclusion of random effect, that is, considering the frailty standard mixture model, we wish to know what is the cost (variance inflation) to estimate p_0 in the presence of α. To realize this study, we proceeded with a study simulation. The generation of the data set and the maximum likelihood estimates were obtained as done in Section 5. The data were generated of the FWM model following sets of parameters $p_0 = 0.5$, $\sigma = 1$,

$\gamma = 1$ and $\alpha = 0.5$, 1, 1.5, 2, 5, $n = 30, 50, 70, 100$ and 300.

Let $\boldsymbol{\vartheta} = (p_0, \sigma, \gamma, \alpha)^\top$ parameters vector of the FWM model and $\widehat{\boldsymbol{\vartheta}} = (\widehat{p}_0, \widehat{\sigma}, \widehat{\gamma}, \widehat{\alpha})^\top$ the maximum likelihood estimates with observed information matrix denoted by J. Let $\widehat{\boldsymbol{\vartheta}}_* = (\widehat{p}_{0*}, \widehat{\sigma}_*, \widehat{\gamma}_*)^\top$ the maximum likelihood estimates of the WM model with observed information matrix denoted by J_*.

The aim is established the estimation cost p_0 in presence of α, that is, measure as the variance of $\widehat{p}_0$ is inflated when α is present in the model. A measure of this cost is given by variances ratio (VR) defined by

$$\mathrm{VR}(\widehat{p}_0) = \frac{J_{11}^{-1}}{J_{*11}^{-1}},$$

where J_{ii} is the ith diagonal element of the observed information matrix J, evaluated in their maximum likelihood estimates, $\widehat{p}_0$, $\widehat{\sigma}$, $\widehat{\gamma}$ and $\widehat{\alpha}$, and J_{*ii} is the ith diagonal element of the observed information matrix J in their maximum likelihood estimates, $\widehat{p}_{0*}$, $\widehat{\sigma}_*$ and $\widehat{\gamma}_*$.

We say there is not estimate cost p_0 in presence of α, if variances ratio $\mathrm{VR}(\widehat{p}_0)$ is equal to 1. In our simulation study a quantity of 1000 samples were generated for each sample sizes. At each sample, we obtained the estimate cost p_0 and later we determined the mean cost. The variances ratios mean are presented in Table 4. The results indicate there are estimate cost when there is a large heterogeneity in the sample (small α), however, as the population becomes more homogeneous (large α), the variances ratios mean assume values close to 1, thus it indicate there is not estimation cost p_0.

Table 4. Variances ratios mean.

	n				
α	30	50	70	100	300
0.5	1.686	1.348	1.122	1.078	1.080
1.0	1.154	1.089	1.085	1.057	1.040
1.5	1.044	1.043	1.030	1.022	1.019
2.0	1.053	1.050	1.017	1.014	1.008
5.0	1.004	1.003	1.002	1.001	1.001

The Figure 1 illustrates for different values of α and n the behavior of mean cost.

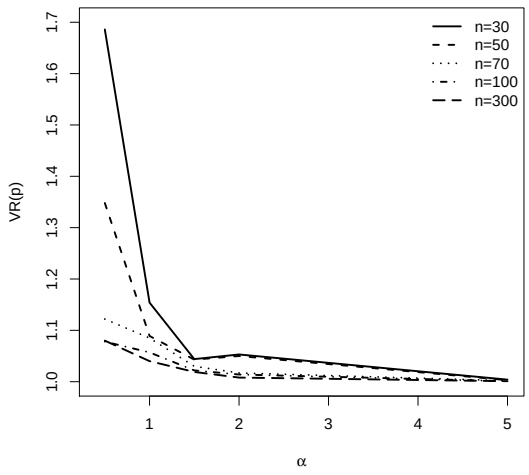

Figure 1. Variances ratios mean.

The simulation study shows that the cost in include the frailty in the model does not inflate significantly the variance of p_0 when α is greater than 2. However, whether there is a large heterogeneity in the population (small α) there is an estimation cost of the cured proportion.

7. Example with leukemia data set

In this section, we attempt to analysis the leukemia data of Kersey et al. (1987). The data set consists of lifetime up to recurrence of leukemia, in years, for a group of 44 patients who received autologous marrow. The observed maximum time was about five years. Censuring is observed for 20% of the lifetimes. The authors reported that the fraction of cured patients was estimated to be 20%.

Before we fit the model, we must verify the shape of the hazard rate function. We follow a standard graphical methodology for data analysis, we use the total time on test (TTT) plot, which is described by Aarset (1985). The empirical version of the TTT plot is given by

$$G(r/n) = \frac{\sum_{i=1}^{r} Y_{i:n} - (n-r)Y_{r:n}}{\sum_{i=1}^{r} Y_{i:n}},$$

where $r = 1, \ldots, n$ and $Y_{i:n}$ $i = 1, \ldots, n$ are the order statistics of the sample.

The increases (decreases) hazard function if the TTT plot has concave (convex) shape. If the TTT plot approximates of a diagonal line we have constant hazard function and, if the curve is concave and then convex, so unimodal hazard function. If the TTT plot has convex shape and then concave shape, we have bathtub shape hazard function. The TTT plot is only a sufficient and not necessary condition to indicate the shape of the hazard function for the individuals at risk and it will be used as an indicator of their behavior. The Figure 2 presents the TTT plot for the leukemia data, implying in increase hazard function. This situation can be represented by Weibull distribution.

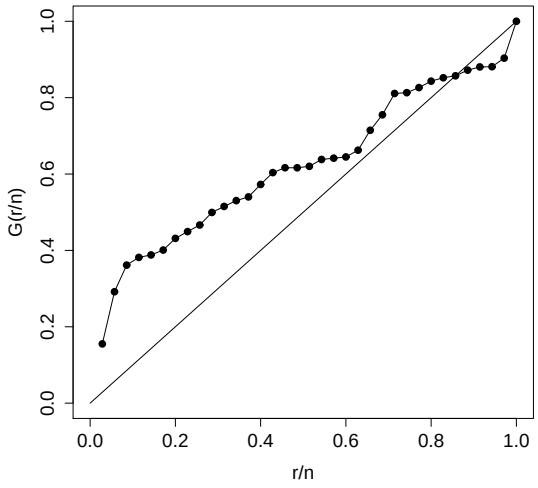

Figure 2. TTT plot to the leukemia data set.

The FMW and WM models are fitted to the data set and the results of maximum likelihood estimates of the parameters and the asymptotic and bootstrap confidence intervals are presented in Table 5. Bootstrap percentile intervals are based on 5000 replicates. The fitted models can be compared employing the log-likelihood function at its maximum, expressed by $\max \mathcal{L}(\cdot)$, the Akaike information criterion (AIC), and the Information Bayesian criterion (BIC). Table 6 presents these statistics from the adjusted models. The FWM model yields the best fitting according to these criteria.

Table 5. Maximum likelihood estimates and approximate confidence intervals from the FWM and WM models.

Model	Parameter	Estimate	95% confidence interval	
			Asymptotic	Bootstrap
FWM	p_0	0.201	(0.083 ; 0.318)	(0.091 ; 0.322)
	σ	6.731	(1.430 ; 12.033)	(4.902 ; 10.412)
	γ	1.862	(1.276 ; 2.447)	(1.587 ; 2.260)
	α	2.535	(0.745 ; 8.637)	(1.011 ; 3.356)
WM	p_0	0.205	(0.104 ; 0.338)	(0.090 ; 0.319)
	σ	3.391	(2.301 ; 4.793)	(2.102 ; 8.015)
	γ	1.371	(1.075 ; 1.684)	(1.111 ; 2.143)

Table 6. Statistics from the adjusted models.

Model	Statistic		
	AIC	BIC	$\max \log \mathcal{L}(\cdot)$
FWM	39.07	45.29	-15.54
WM	44.88	49.55	-19.44

The results presented in Table 5 shows that the estimation cure rate in both models are close, this fact was expected as seen in the simulation study. We also note that the asymptotic and bootstrap confidence intervals are close. On the other hand, the bootstrap confidence intervals for the parameters σ, γ and α have lower amplitude. For the WM model, the bootstrap confidence intervals show upper amplitude than asymptotic intervals.

The Figures 3 and 4 show the wormplot of the quantile residuals for the FWM and WM models, respectively.

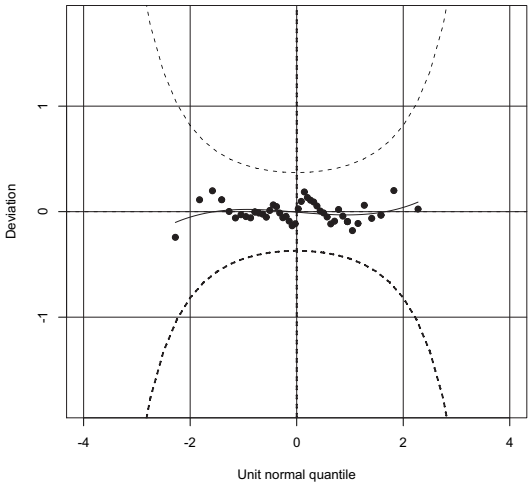

Figure 3. Wormplot of the quantile residuals for the FWM model.

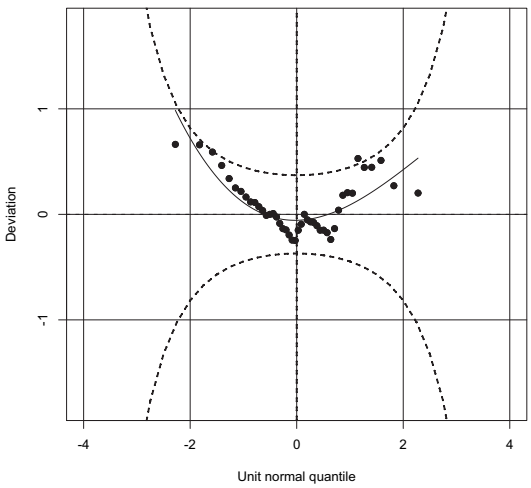

Figure 4. Wormplot of the quantile residuals for the WM model.

The wormplot indicate the FWM model as best fitted, because all the points belong to the region bounded by the two dashed curves. In Figure we see that for the WM model there are points not belonging to this central region, indicating that this model is not appropriate for this data set.

In the presence of long-term survival data, we are interested in estimating the cured rate of the population, so it is reasonable to proceed with a test to verify if there are really immune individuals in the population, that is, we are interested in testing the null hypothesis $\mathrm{H}_0 : p_0 = 0$. The statistic test most used is likelihood ratio. Asymptotically this statistic has distribution χ^2_1, but under H_0 the parameter value is on the boundary of the parametric space and problems can occur when testing the null hypothesis mentioned.

Let $\boldsymbol{\theta}$ denote the parameter vector of lifetime distribution and $\mathcal{D}$ is the observed data. The likelihood ratio test (LRT) is given by

$$\Lambda = 2\left(\ell\left(\widehat{p_0}, \widehat{\alpha}, \widehat{\boldsymbol{\theta}}\,; \mathcal{D}\right) - \ell\left(0, \widetilde{\alpha}, \widetilde{\boldsymbol{\theta}}\,; \mathcal{D}\right)\right), \tag{10}$$

where $\widetilde{\alpha}$ and $\widetilde{\boldsymbol{\theta}}$ are the maximum likelihood estimates under null hypothesis.

Under certain regularity conditions, Maller and Zhou (1995) showed that the statistical distribution Λ is a mixture in proportions $(50\%, 50\%)$ of a chi-square distribution with one degree of freedom and a point mass at zero, that is

$$\mathbb{P}(\Lambda \leq \xi) = \frac{1}{2} + \frac{1}{2}\mathbb{P}(\chi^2_1 \leq \xi). \tag{11}$$

The 95th percentil of distribution given in Equation (11), it is represented by $\xi_{0.95}$, is such that

$$\frac{1}{2} + \frac{1}{2}\mathbb{P}\left(\chi^2_1 \leq \xi_{0.95}\right) = 0.95,$$

so $\xi_{0.95} = 2.71$. Therefore, we reject H_0 significance at 5% level if $\Lambda > 2.71$.

For this, we proceeded with the hypothesis test to check if there are immune individuals in the population. From Equation (10), we obtain $\Lambda = 35.18 > 2.71$. Therefore, we have evidence of immune individuals in the population. To check if the unobservable heterogeneity is present, we used the LRT considering the usual gamma frailty model and the

homogeneous survival model, obtaining $\Lambda = 37.06$, that is, there is evidence of a degree of unobservable heterogeneity in the sample.

This way, we must include frailty in the standard mixture model. Considering the FWM model, we need to check if there are evidence of immune individuals. We realized the hypothesis test and we obtained $\Lambda = 10.15 > 2.71$, that is, there are evidence of cured individuals in population.

Taking into account the results in Table 6, Figures 3 and 4 and hypothesis tests, we selection the FWM model as our working model. All parameters are significant at a 5% level. The cured rate was estimated in 0.201, the estimates of the scale and shape parameters of the Weibull distribution are, respectively, $\sigma = 6.731$ and $\gamma = 1.862$, what means increase failure rate. With the inclusion of frailty term, we can quantify the amount of heterogeneity among individuals and for this example the unobservable heterogeneity is $1/\alpha = 0.395$.

The Figure 5 presents the Kaplan-Meier curve jointly with the estimated survival function for the FWM and WM models.

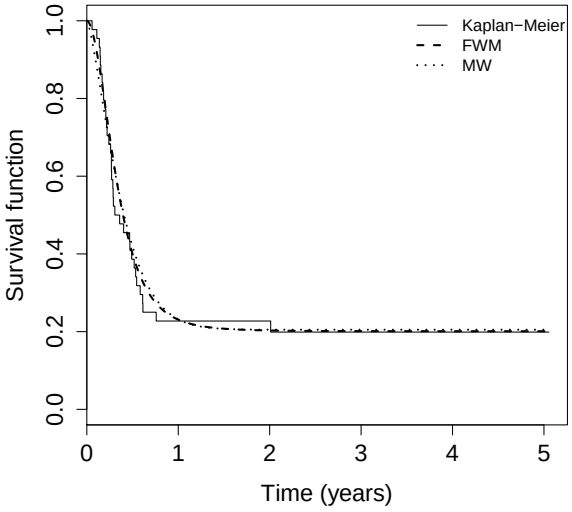

Figure 5. Survival functions of the fitted models jointly Kaplan-Meier curve.

8. Mode of availability

Computational codes of the long-term survival models and the example in Sections 5, 6 and 7 can be downloaded from *http://www.ufscar.br/jcfogo/Files/Programs.rar*.

9. Conclusions

The standard mixture model by Berkson & Gage implicitly assume for the individuals susceptible homogeneous risk, however, we know that the individuals are different and consequently there are biological variations among those individuals that can not be measure. With aim to quantify the amount of heterogeneity among subjects, we add a frailty term and we suppose a Weibull distribution for the lifetime, obtaining the FWM model.

With the studies realized in this paper, we can conclude that realize a traditional analysis when there is a degree of unobservable heterogeneity the parameter estimates of the

Weibull distribution are significantly affected, however, for the cured fraction there is not problem. Through of simulations study, we quantify this error, which we compare the mean of estimates of the models, the standard error and mean squared error. For the frailty parameter α, the results showed that for small sample there are biased estimates. Therefore, asymptotic results should be used with caution. From a practical standpoint, the frailty standard mixture model provide better fit to the leukemia data set, since it allowed to quantify the amount of unobservable heterogeneity among individuals beyond of estimate of the cured fraction. Although further research must be conducted in regard to the models, initial investigation suggests will serve as an enhancement to the field of survival analysis.

Acknowledgements

The authors are grateful for the fruitful comments and suggestions of two anonymous referees.

References

Aalen, O., 1988. Heterogeneity in survival analysis. Statistics in Medicine 7, 1121–1137.

Aarset, M., 1985. The null distribution for a test of constant versus "bathtub" failure rate. Scandinavian Journal of Statistics 12, 55–61.

Andersen, P., 1993. Statistical models based on counting processes. Springer Verlag.

Anscombe, F., 1961. Estimating a mixed-exponential response law. Journal of the American Statistical Association 56, 493–502.

Berkson, J., Gage, R., 1952. Survival curve for cancer patients following treatment. Journal of the American Statistical Association 47, 501–515.

Boag, J., 1949. Maximum likelihood estimates of the proportion of patients cured by cancer therapy. Journal of the Royal Statistical Society B 11, 15–53.

Broadhurst, R., Maller, R., 1991. Estimating the in criminal careers from one-step probabilities of recidivism. Journal of Quantitative Criminology 7, 275–290.

Clayton, D., 1978. A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. Biometrika 65, 141–151.

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

Elbers, C., Ridder, G., 1982. True and spurious duration dependence: The identifiability of the proportional hazard model. The Review of Economic Studies 49, 403–409.

Farewell, V., 1977. A model for a binary variable with time-censored observations. Biometrika 64, 43–46.

Goldman, A., 1984. Survivorship analysis when cure is a possibility: a Monte Carlo study. Statistics in Medicine 3, 153–163.

Hougaard, P., 1995. Frailty models for survival data. Lifetime Data Analysis 1, 255–273.

Hougaard, P., Myglegaard, P., Borch-Johnsen, K., 1994. Heterogeneity models of disease susceptibility, with application to diabetic nephropathy. Biometrics 50, 1178–1188.

Kersey, J., Weisdorf, D., Nesbit, M., LeBien, T., Woods, W., McGlave, P., Kim, T., Vallera, D., Goldman, A., Bostrom, B., et al., 1987. Comparison of autologous and allogeneic bone marrow transplantation for treatment of high-risk refractory acute lymphoblastic leukemia. New England Journal of Medicine 317, 461–467.

Longini Jr, I., Halloran, M., 1996. A frailty mixture model for estimating vaccine efficacy. Applied Statistics 45, 165–173.

Maller, R., Zhou, S., 1995. Testing for the presence of immune or cured individuals in censored survival data. Biometrics 51, 1197–1205.

Maller, R. A., Zhou, X., 1996. Survival analysis with long-term survivors. Wiley, New York, NY.

Meeker, W., Escobar, L., 1998. Statistical methods for reliability data. Wiley, New York, NY.

Oakes, D., 1982. A model for association in bivariate survival data. Journal of the Royal Statistical Society B 44, 414–422.

Peng, Y., Taylor, J., Yu, B., 2007. A marginal regression model for multivariate failure time data with a surviving fraction. Lifetime Data Analysis 13, 351–369.

Price, D., Manatunga, A., 2001. Modelling survival data with a cured fraction using frailty models. Statistics in Medicine 20, 1515–1527.

R Development Core Team, 2012. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Rigby, R., Stasinopoulos, D., 2005. Generalized additive models for location, scale and shape. Journal of the Royal Statistical Society: Series C (Applied Statistics) 54, 507–554.

Rodrigues, J., Cancho, V., de Castro, M., Louzada-Neto, F., 2009. On the unification of long-term survival models. Statistics & Probability Letters 79, 753–759.

Shao, Q., Zhou, X., 2004. A new parametric model for survival data with long-term survivors. Statistics in Medicine 23, 3525–3543.

Sinha, D., Dey, D., 1997. Semiparametric Bayesian analysis of survival data. Journal of the American Statistical Association 92, 1195–1212.

Stasinopoulos, D., Rigby, R., 2007. Generalized additive models for location scale and shape (GAMLSS) in R. Journal of Statistical Software 23, 1–46.

Vaupel, J., Manton, K., Stallard, E., 1979. The impact of heterogeneity in individual frailty on the dynamics of mortality. Demography 16, 439–454.

Wienke, A., 2011. Frailty Models in Survival Analysis. Chapman & Hall/CRC Biostatistics Series. CRC Press.

Yau, K., Ng, A., 2001. Long-term survivor mixture model with random effects: application to a multi-centre clinical trial of carcinoma. Statistics in Medicine 20, 1591–1607.

Yu, B., Peng, Y., 2008. Mixture cure models for multivariate survival data. Computational Statistics & Data Analysis 52, 1524–1532.