"Science Stays True Here"
Journal of Mathematics and Statistical Science, 179-198 | Science Signpost Publishing

Frailty Modeling for Repairable Systems with Minimum Repair: An Application to Dump Truck Data of a Brazilian Mining Company

Amanda Morales Eudes D'Andrea, Cirdêmia Costa Feitosa

Universidade Federal de São Carlos, Departamento de Estatística, São Carlos, SP, Brasil e Universidade de São Paulo, Instituto de Ciências Matemáticas e de Computação, São Carlos, SP, Brasil.

Vera Lucia Damasceno Tomazella, Afrânio Márcio Corrêa Vieira

Universidade Federal de São Carlos, Departamento de Estatística, São Carlos, SP, Brasil.

Abstract

In repairable system data analysis it is common to many components of the same type are studied and in these cases it is relevant to verify the heterogeneity between systems. Proschan [1] pointed out that the unobserved heterogeneity may explain increasing failure rates, which is often found in reliability analysis. The unobserved heterogeneity may be estimated from models of frailty. This model is characterized by using a random effect, that is, a non-observable random variable that represents the information that could not or were not observed. However, according with Vaupel et al. [2], the standard methods in repairable system data analysis ignore the unobserved heterogeneity. Thus, this work will explore the frailty models. The inferential method for estimation of the parameters will be displayed for models with minimal repair. Finally, an application to real data set [3] was taken, in which the models with and without frailty are compared. These real data set involving failures in trucks was collected in a Brazilian mining company.

Keywords: *Repairable systems, minimal repair, Power Law Process, frailty*.

1. Introduction

One of the major problems of the activities carried out in series is the occurrence of a failure in some equipment at any stage of the process. For example, faults on mats in manufacturing engineering

Corresponding author: Vera Lucia Damasceno Tomazella, Universidade Federal de São Carlos, Departamento de Estatística, São Carlos, SP, Brasil.

companies, failures in logistics companies to transport vehicles, failures in software components, damage in complex equipment such as control panels or problems in carriers of truck fleets.

The occurrence of failure in a machine can generate for the company costs that exceed the budget, delays in the delivery of their products or services and to loss of goods and customers. So occurred unwanted and inherent failure, it is essential and crucial for the resumption of the process initially planned that a repair action to take place as soon as possible and in the best possible way.

A definition commonly used in the literature to define repairable systems was given by Ascher and Feingold [4], that a system, whether machinery, electronic equipment or software, is considered repairable if after a failure, its activity can be satisfactorily resumed through repair without the need to replace the system as a whole. The authors reported that researchers and reliability professionals did not recognize the crucial difference between repairable and non-repairable systems and demonstrated with several simple examples as wrong conclusions can be taken if this difference is not considered. As noted by Lindqvist [5], the book Ascher and Feingold [4] seems to be the first dedicated exclusively to the reliability of repairable systems and for a long time this was the main reference in the literature and is still a major source of quote.

When considering models for repairable systems, a critical point is how to account for the effect of repair actions taken after failures. In this sense, the most explored assumptions are Minimal Repair (MR), which returns the system to the condition just before the failure (ABAO - As Bad as Old), and Perfect Repair (PR), which leaves the system as if it were new (AGAN - As Good as New). These assumptions were discussed in works as Phelps [6], Barlow and Proschan [7], Zhao and Xie [8], Park et al. [9] and Wang [10].

The minimal repair (MR) is the assumption further explored in the literature that focuses on correct only the component originator of failure, leaving the system in the same condition it was prior to it. Here, the process associated with the occurrence of failure can be described by a Non-Homogenous Poisson process (NHPP), in which the probability of failure in a short time depends not on the fault history, only the system’s age, [11]. According Kijima [12], this assumption is plausible for systems consisting of several components, each having its own failure mode. These models are based on counting process. Aalen [13] was the first author to develop statistical models based on counting processes for recurrent events. Tomazella [14] reports on the extensive literature on the models of counting processes and that the models often used for recurrent event data that provide learning on the individual case, are those based on

Poisson process and renewal process known as intensity patterns. Some examples where covariates are not present in these models are Cox and Isham [15] and Cox and Lewis [16].

In repairable system data analysis it is common withdraw many components of the same type are studied and in these cases it is relevant to verify the heterogeneity between systems. Proschan [1] pointed out that the unobserved heterogeneity may explain increasing failure rates, which is often found in reliability analysis. The heterogeneity observed may be estimated from models of frailty. This model is characterized by using a random effect, that is, a non-observable random variable that represents the information that could not or were not observed. According Vaupel et al. [2], the standard methods in repairable system data analysis ignore the unobserved heterogeneity. Consider that the systems have different frailties and those who are most fragile fail earlier than others is the approach of this study.

The work was motivated by a study of Toledo [3] related to maintenance problems in a Brazilian mining trucks. Large trucks, known as off-road trucks, transporting the ore from the mine to the homogenization pile. For the supply of ore to the treatment plant is constant, if these trucks off road show failure, dump trucks are put into operation, while the maintenance is done in the other.

The dump trucks were not designed to operate under the conditions found at mining sites, but help in the process as a whole. When these trucks are flawed, often their repair is immediate, because its purpose is to provide assistance to large trucks. In the study by Toledo [3] with dump trucks, the failure times between the different trucks were considered independent. However, it is important to check whether there is heterogeneity between systems because there may be information such as environmental conditions, stress, use, among others, which were could not or not included in the study but that somehow are important for their analysis.

Thus, the purpose of this paper is to approach the minimal repair model and consider risk factors unobservable to model and characterize the unobservable heterogeneity between systems by adding a term of frailty to the intensity function of the minimal repair model, which assumes a Gamma distribution for the variable frailty. A versatile parametric form and extremely explored in repairable systems is the Power Law Process (PLP), and this will be considered in this work.

This paper is organized as follows. In Section 2, a brief introduction to the usual models for repairable systems is made, focusing on minimal repair model and the estimation method of the parameters of the proposed model. In the following, some concepts about the frailty model are presented. In Section 3, is presented the insertion of a random effect on the intensity function of the model repairable systems. A simulation study were performed in Section 4 and an application to a real data set in Section 5.

The work is completed in Section 6 with final considerations appears.

2. Background

2.1 Minimal Repair Models

Typical models for repairable systems are the minimal, perfect and imperfect repair. The type of repair to be done depends on the type of system and of course the type of fault displayed. Using the minimal repair concept, it is possible to describe in a simple way the fact that many repairs in real life bring the system to a condition which is basically the same as it was just before the failure occurred. Such a repair may be used to model a system where a component of the system is replaced or repaired. Of course, the purpose of the repair action is not to bring the system to the exact same condition. Rather the purpose is to bring the system back to operation as soon as possible. But by looking at the condition of the system after the repair, it is a reasonable assumption to say that the system state has not changed.

The minimal repair (MR) focuses on correct only the component originator of failure, leaving the system in the same condition it was in prior to it, known in the literature by *As Bad as Old - ABAO*. The process associated with the occurrence of failure can be described by a Non-Homogenous Poisson process (NHPP), in which the probability of failure in a short time not dependent on its fault history, but only the old system [11] and thus, the intensity function will depend only on the time. According Kijima [12], this assumption is plausible for systems consisting of several components, each having its own failure mode.

This means that the conditional intensity of the failure process immediately after the failure is the same as it was immediately before the failure, and hence is exactly as it would be if no failure had ever occurred. Thus a process $N(t)$ with intensity function $\lambda(t)$ is given by

$$\rho_{MR}(t) = \lambda(t)$$

and the process is an NHPP. In practice a minimal repair usually corresponds to repairing or replacing only a minor part of the system.

The minimal repair is the most exploited assumption in the literature and many special cases of the basic minimal repair model have been addressed in the literature. Two of the most frequently studied special cases are Power law Process and Log linear model.

Be $N(t)$ the number of failures from the beginning of system monitoring until a certain time t. if $N(t)$ follows a Power Law Process (PLP), the intensity function, and cumulative intensity function proposed by Crow [17] are given respectively by

$$\lambda(t) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} \quad (1)$$

$$\Lambda(t) = \left(\frac{t}{\eta}\right)^{\beta} \quad (2)$$

where $\eta > 0$ is the scale parameter and $\beta > 0$ is the shape parameter.

The parameter η can be interpreted as the time during which expected exactly a failure occurs, that is, $E[N(\eta)] = 1$, while β is the elasticity of the expected number of failures with respect to time, so when $\beta > 1$, the system is deteriorating (see Oliveira et al. [18]).

According to Crow [17], this process is common in the literature of repairable systems and Oliveira et al. [18] mentions that its popularity is due to the fact that their function be flexible. More precisely, the PLP can accommodate both intensities of increasing occurrences which occur when $\beta > 1$, as decreasing intensities $\beta < 1$, or when occurrence intensity of the event of interest is constant for where $\beta = 1$, featuring a Homogenous Poisson process (HPP). Therefore, the PLP is characterized in that this spring is constant over time.

Thus, the intensity function for the MR is being equal to the 1. This means that the intensity immediately after the failure process is the same found immediately before the failure and then, is exactly how it could be if the failure had not occurred, [5]. In practice a minimal repair usually corresponds repair or replace only a minor part of the system.

To perform the estimation of the parameters, the method used is the maximum likelihood. The necessary functions are presented as follows.

2.1.1 Inference

In repairable systems there are basically two ways to look at the data. At first, data collection ends at a pre-set time t, called by truncation by time, known in analyzing reliability/survival as censorship Type II. In the second, data collection ends after the occurrence of a set number of failures, denoted by truncation for failure analysis and reliability/survival, this type of study is known as Type censorship I. In both truncation, the censures will always appear at the end of the study.

Consider k repairable systems, for $k = 1, 2, \ldots$, where the systems are independent, assuming the following conditions:

- The *i*-th truncated system time is observed until the time pre-determined t_i^*;
- n_i failures are observed in the *i*-th system, truncated by time or failure, $i = 1, 2, \ldots, k$;
- $N = \sum_{i=1}^{k} n_i$ is the total number of observed failures in the systems;
- Let $t_{i,j}$, $i = (1, 2, \ldots, k)$ and $j = (1, 2, \ldots, n_i)$, observations of the random variable T that represent the failure times for the *i*-th system, recorded as the time from the start of the system $(t_{i,1} < t_{i,2} < \ldots < t_{i,n_i})$;
- Let $\boldsymbol{\mu} = (\beta, \eta)$ vector parameters to be estimated.

The likelihood function for this process should combine the joint probability density of the failure times of the k systems and using the intensity functions under the minimal repair assumption and the parametric form of the Power Law Process.

When the study is done by truncation of time, the likelihood function under the MR model assuming PLP is given by

$$L_{MR}(\boldsymbol{\mu}) = \prod_{i=1}^{k} \left\{ \prod_{j=1}^{n_i} \left[\frac{\beta}{\eta} \left(\frac{t_{i,j}}{\eta} \right)^{\beta-1} \right] \exp\left[-\left(\frac{t_i^*}{\eta} \right)^{\beta} \right] \right\} \tag{3}$$

and the logarithm of the likelihood function is given by

$$\ell_{MR}(\boldsymbol{\mu}) = \sum_{i=1}^{k} \left\{ \sum_{j=1}^{n_i} \log\left[\frac{\beta}{\eta} \left(\frac{t_{i,j}}{\eta} \right)^{\beta-1} \right] \right\} + \sum_{i=1}^{k} \left[-\left(\frac{t_i^*}{\eta} \right)^{\beta} \right]. \tag{4}$$

In cases where the study was done by truncation of failure, for obtain the likelihood function just replace $t_i^* = t_{i,n_i}$.

When solving the estimating equations based on (4), it is not possible to obtain analytical solution and numerical methods are considered for estimation of parameters.

2.2 Frailty models

In the literature, relating to the various models of frailty, authors studied the use of fragile multiplicative models, and they represent an extension of Cox [19]. From the classical point of view, Andersen et al. [20] and Hougaard [21] submitted a review of these models, whereas Sinha and Dey [22] prepared a review under the Bayesian approach. In univariate data, the term frailty was inserted by Vaupel et al. [2], Clayton [23] and Oakes [24] work with models for multivariate cases.

The frailty multiplicative model introduces a random effect, called frailty, in the intensity function for the purpose of describing the possible association between the units in order to control the unobservable heterogeneity of the studied systems. Thus, the risk of a system depends on a non-observable random variable, non-negative, which acts multiplicative in the intensity function.

Thus, the failure intensity for the system i, for $i = 1, \ldots, k$, is given by

$$\rho_i(t \mid z_i) = z_i \lambda_0(t_{i,j}), \tag{5}$$

of which $\lambda_0(t_{i,j})$ ($j = 1, \ldots, n_i$) is the common intensity function to systems and z_i is the variable of frailty independent and identically distributed with function density $P(z_i)$.

The conditional reliability function is established by

$$R_i(t \mid z_i) = [R_0(t_{i,j})]^{z_i} = exp[-\Lambda_0(t_i^*) z_i], \tag{6}$$

where z_i is the frailty of the *i*-th system, $\Lambda_0(t)$ is the common cumulative intensity function to all units and t_i^* is the truncation of time, but if the study is truncated by failure, simply $t_i^* = t_{i,n_i^*}$.

As z_i represents a value of the random variable unobservable Z when $z_i > 1$ means that the individual risk is increasing and when $z_i < 1$ the risk is decreasing, already for $z_i = 1$ model frailty (5) is reduced to the common intensity function to units. In this model, as the frailty variable acts multiplicative, it follows that the larger the value of this variable, the greater the chances of the occurrence of failure, that is, the greater the value of z_i, more "fragile" units will be, hence the frailty name. Therefore, for these weaker units it is expected that the event of interest occurs more often.

In general, it is assumed that the frailties are independent and identically distributed for each unit. Assuming this term of frailty follows a probability distribution, the choice of such distribution is an important point to be addressed. Because of this frailty term act in multiplicative manner in intensity function, the candidates to the distribution of frailty are supposed not negative, usually continuous and not dependent on time, such as gamma distribution, log-normal, inverse Gaussian and Weibull. Because of the fact the gamma distribution is easy algebraic treatment, it has been widely used to model the frailty. The first of frailty model for survival data using multivariate distribution range was considered by Clayton [23] and Oakes [24].

Hougaard [25] suggested for the frailty variable inverse Gaussian distribution as an alternative to gamma distribution, since the distribution becomes more homogeneous population over time, which is compatible with the idea that the event of interest occurs primarily to the weaker units. The gamma and inverse Gaussian distributions satisfy assumptions of identifiability, in which the frailty Z is a non-negative random variable with distribution function $P(Z)$ and as described in [26]. From the computational point of view, the distribution range fits very well to the reliability models, because it is easy to derive formulas for some event numbers.

In this work it will be considered the distribution range, so, assuming that the frailty variable has Gamma distribution (α,α), the probability density function for Z is given by

$$f(z) = \frac{\alpha^{\alpha}}{\Gamma(\alpha)} z^{\alpha-1} e^{-\alpha z}, \tag{7}$$

to expected value $E(Z)=1$, thus avoiding the problems of identifiabity, and variance $V(Z)=1/\alpha$, and such variance is the dependent parameter.

The complete likelihood function that contains the observed data ($t_{i,j}$) not observed (z_i) is given by (see Lawless [27])

$$L(\boldsymbol{\mu} \mid z_i, t_{i,j}) = \prod_{i=1}^{k}\prod_{j=1}^{n_i} \rho(t_{i,j} \mid z_i) R(t_{i,j} \mid z_i), \tag{8}$$

where the random variable z_i has distribution $f(z)$, $\boldsymbol{\mu}$ is the parameter vector and n_i is the total number of failures observed in the *i*-th system. Then, when seeking distributions for the frailty variable Z, it is natural to use frailty distributions with an explicit Laplace transform, because it facilitates the use of standard ML methods for parameter estimation.

To obtain the unconditional reliability function we need to integrate out the frailty term, that is,

$$R(t) = \int_0^{\infty} R(t \mid z) f(z) dz. \tag{9}$$

In order to obtaining the unconditional hazard we start from the unconditional reliability and take negative logs to obtain the cumulative hazard given by

$$H(t) = -\log R(t) \tag{10}$$

$$= -\log \int_0^\infty R(t \mid z) f(z) dz. \tag{11}$$

A very useful tool in frailty analysis is the *Laplace transform* (see Wienke [28]). Given a function $f(x)$, the Laplace transform considered as function of a real argument s is defined as,

$$Q(s) = \int_0^\infty e^{sx} f(x) dx. \tag{12}$$

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, $f(z)$ and s as the cumulative baseline hazard $H_0(t)$ and we obtain

$$R(t) = \int_0^\infty e^{H_0(t)z} f(z) dz = Q(H_0(t)), \tag{13}$$

where Q denotes the Laplace transform.

The frailty random variable Z_i are usually assumed to be independent with identical frailty distribution. As mentioned, the frailty distribution can be gamma, Inverse Gaussian, Log Normal or Weibull. This work we consider the Gamma distribution.

3. Frailty Model under the Assumption of Minimal Repair

The main objective of this work is considered the frailty variable for models of repairable systems. It was seen in the section 2.2 how to find the probability density function non conditioned to frailty. Following will be presented the intensity functions and unconditioned risk frailty to the minimal repair model.

Consider the intensity function of frailty model (5), whose base function $\lambda_0(t)$ is the function (PLP). This way you can set the model of interest, minimal repair model with frailty, whose intensity conditioned MR function the frailty variable is given by

$$\rho_i(t \mid z_i) = z_i \lambda_0(t) = z_i \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1}. \tag{14}$$

The MR reliability function conditioned to the variable $Z = z$ is established by

$$R_i(t \mid z_i) = [R_0(t)]^{z_i} = \left\{ \exp\left[-\left(\frac{t^*}{\eta} \right)^{\beta} \right] \right\}^{z_i} \tag{15}$$

where z_i represents the frailty of i th system.

Considering the likelihood function of the MR model (3) and the insertion of the multiplicative frailty term the way it was made in frailty model (5) and function (6) one can rewrite function (8) as follows

$$L(\boldsymbol{\mu}, z_i) = \prod_{i=1}^{k} \prod_{j=1}^{n_i} \left[z_i \lambda_0(t_{i,j}) \right] \mathrm{e}^{-\Lambda_0(t_i^*) z_i} f(z). \tag{16}$$

Integrating (16) compared to z_i , and $f(z_i)$ is (7), the marginal partial likelihood function, being n_i failures observed by times $t_{i,j}$ ($j = 1, \ldots, n_i$) to the system i is given by

$$L_i = \int_0^{\infty} \left[\prod_{j=1}^{n_i} z_i \lambda_0(t_{i,j}) \right] \mathrm{e}^{-\Lambda_0(t_i^*) z_i} \frac{\alpha^{\alpha}}{\Gamma(\alpha)} z_i^{\alpha-1} \mathrm{e}^{-\alpha z_i} dz_i. \tag{17}$$

To determine the marginal likelihood function, simply obtain the non conditional to frailty reliability function that, considering the function (6), it is needed to integrate the term of frailty as follows

$$R(t) = \int_0^{\infty} R_i(t \mid z_i) f(z) dz = \int_0^{\infty} \left[R_0(t) \right]^{z_i} f(z) dz = \int_0^{\infty} [\mathrm{e}^{-\Lambda_o(t^*)}]^{z_i} f(z) dz, \tag{18}$$

where $f(z)$ is the probability density function (pdf) of the frailty variable and $\Lambda_0(t)$ is the cumulative intensity function of $\lambda_0(t)$.

Considering that $f(z)$ has gamma distribution with pdf (7), the Laplace transformation of the Gamma (α, α) distribution is given by

$$L_g[s] = \left(1 + s\alpha^{-1} \right)^{-\alpha}. \tag{19}$$

Thus, in the function (19) simply $s = \Lambda_0(t)$ and, in this way, the conditional to frailty reliability function is determined by

$$R(t) = \left(1 + \Lambda_0(t) \alpha^{-1} \right)^{-\alpha}. \tag{20}$$

And then, the non conditional to frailty intensity function is given by

$$\rho(t)=\lambda_0(t)\left(1+\Lambda_0(t)\alpha^{-1}\right)^{-1}. \tag{21}$$

Therefore, being $t_{i,j}$ the *j*-th failure time of the *i*-th system, from the functions (20) and (21) obtained from the Laplace transformation, the partial likelihood function considering the MR model (3) is determined by

$$L(\boldsymbol{\mu}\,|\,t_{i,j})=\prod_{i=1}^{k}\prod_{j=1}^{n_i}\left[\lambda_0(t_{i,j})\right]\left[\left(\Lambda_0(t_i^*)\alpha^{-1}+1\right)\right]^{-(n_i+\alpha)} \tag{22}$$

$$=\prod_{i=1}^{k}\prod_{j=1}^{n_i}\left[\frac{\beta}{\eta}\left(\frac{t_{i,j}}{\eta}\right)^{\beta-1}\right]\left(\left(\frac{t^*}{\eta}\right)^{\beta}\frac{1}{\alpha}+1\right)^{-(n_i+\alpha)}, \tag{23}$$

where t_i^* is the time truncation, but if the study is with failure truncation, make $t_i^*=t_{i,n_i^*}$, according describe in the Section (2.1.1).

The logarithm of the likelihood function (23) is given by

$$\ell_{MR}(\boldsymbol{\mu}\,|\,t_{i,j})=N\log\left(\frac{\beta}{\eta}\right)+(\beta-1)\sum_{i=1}^{k}\sum_{j=1}^{n_i}\log\left(\frac{t_{i,j}}{\eta}\right)-\sum_{i=1}^{k}(n_i+\alpha)\log\left(\left(\frac{t_i^*}{\eta}\right)^{\beta}\frac{1}{\alpha}+1\right),$$

where k is the total number of systems in study and $N=\sum_{i=1}^{k}n_i$ is the total number of occurred failures. As it is not possible to obtain an analytical solution for the estimating equations, numerical methods are considered for estimation of parameters.

4. Simulation Study

In the simulation process performed, first, the parameters were set at $\beta=1.1$, $\eta=6$ and $\alpha=43$, that represents a variability of $\frac{1}{43}=0.023$, considering a type of truncated by failure study, in this case it was considered that the k systems have 5 failures.

We used three sample sizes: $k=30$, $k=50$, $k=150$ and $k=300$ and were generated $d=1000$ replicas. To meet the objectives of the simulation study, we used the following algorithm:

1) Set parameter values;
2) Generate $z_i \sim Gamma(\alpha,\alpha)$;

3) Generate $u_i \sim U(0,1)$;

4) Generate $v_{i,j}$, such that $v_{i,j} = \left(\frac{-\log(1-u_i)}{z_i}\right)^{\frac{1}{\beta}} \eta$;

5) Repeat steps 2 to 5 k times;

6) The simulated times will be $t_{i,j} = t_{i,j-1} + v_{i,j}$, for $j = 2,3,4,5$ and $t_{i,j} = v_{i,j}$ for $j = 1$.

7) Repeat steps 2 to 6 until d replicas.

Each simulation was obtained the mean square error (MSE) for each estimate, the standard error (SE), the amplitude and the 95% confidence interval (CI), based on asymptotic theory.

The results are summarized in Table (1) and Figure (1), where MLE is the maximum likelihood estimator. The results show that as the sample size increases, the amplitude of the confidence interval, the mean square error and the average deviation decreases, as is expected in a simulated for maximum likelihood estimators. Moreover, it is seen that the average of the estimated values for the three parameters, β, η and α approximate the values set as the sample size increases, according to the desired MLE.

Table 1: Results of the study of simulation of the MR model with frailty

Systems	Parameter	MLE	SE	MSE	CI (95%)	Amplitude
30	β	1.0830	0.0827	0.0072	[0.9210 ; 1.2450]	0.3240
	η	6.5690	0.9415	1.2887	[4.7237 ; 8.4143]	3.6906
	α	30.8648	7.3527	169.6789	[16.4538 ; 45.2758]	28.8220
50	β	1.0696	0.0632	0.0049	[0.9457 ; 1.1934]	0.2476
	η	6.4390	0.7237	0.7171	[5.0205 ; 7.8575]	2.8370
	α	33.1918	6.3184	113.8434	[20.8081 ; 45.5756]	24.7675
150	β	1.0637	0.0362	0.0027	[0.9927 ; 1.1347]	0.1419
	η	6.4059	0.4171	0.3453	[5.5884 ; 7.2234]	1.6350
	α	38.9120	4.5889	28.7822	[29.9179 ; 47.9061]	17.9882
300	β	1.0633	0.0256	0.0020	[1.0132 ; 1.1135]	0.1003
	η	6.4150	0.2951	0.2652	[5.8367 ; 6.9933]	1.1566
	α	42.8116	3.7268	8.1625	[35.5072 ; 50.1160]	14.6088

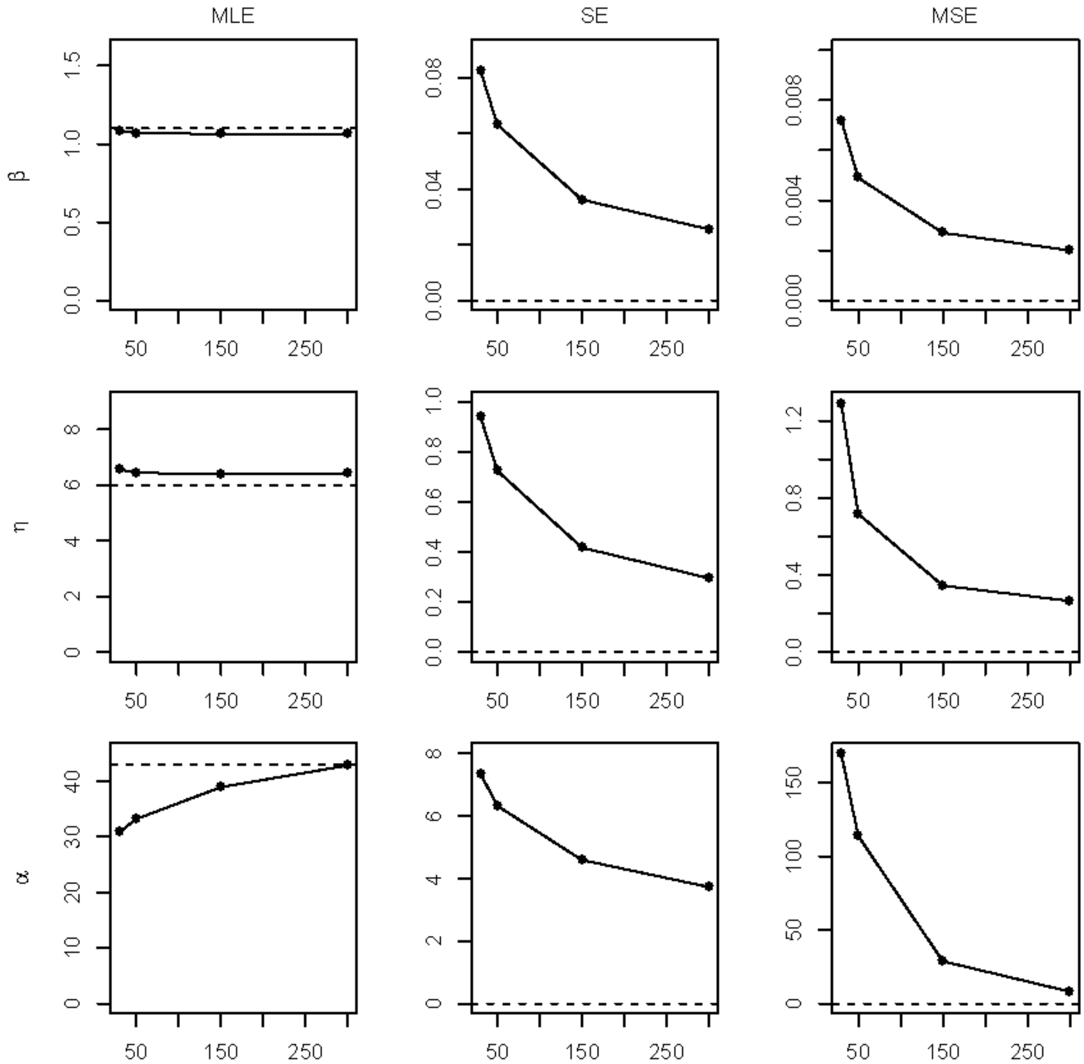
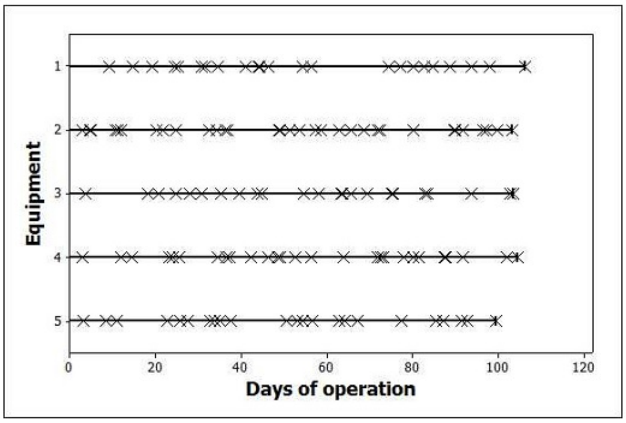


Figure 1: Graph simulation results for the MR model with frailty

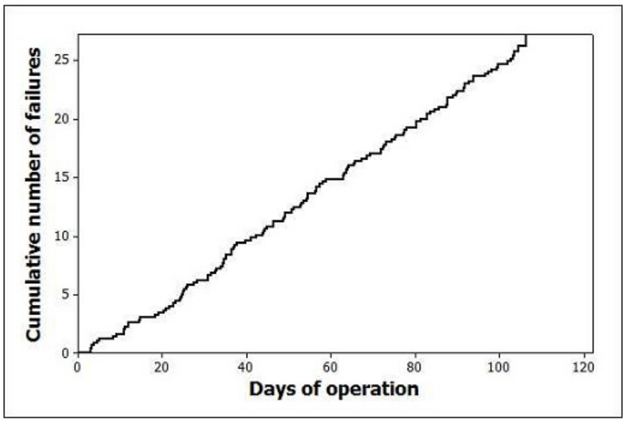
5. Application to Dump Trucks Data

In this section, we illustrate the proposed methodology to quantify the heterogeneity among dump trucks from Brazilian mining company, described in the Introduction. The goal here is to study which model better fits the data, considering the minimal repair model with and without frailty.

Figure (2) presents descriptive plots for the data set, which consists of failure records in a sample of five trucks from the mining company fleet. Data were collected from July to October 2012, when 129 failures were observed, each one followed by a repair. The data for the five trucks were failure truncated, what means the last observation for each one corresponds to a failure time. Figure (2a) shows events (failures) vs operation time (in days), where each line corresponds to a sample unit, and each "x" symbol represents a failure time. Visually, no trend in failures over time can be observed from this graph. Figure (2b) exhibits the mean observed cumulative number of failures. Globally, this curve is neither concave nor convex, so neither improvement nor degradation of the observed equipment can be detected through this visual inspection.



(a)



(b)

Figure 2: (a) Failure times in days of operation for each truck (horizontal lines are trucks and "x" are failures); (b) Cumulative number of failures versus days of operation.

It was used the R software to obtain the estimates of the MR model with frailty parameters. The results follow in Table (2), wherein the estimate for $\hat{\beta} = 1.139$ and it is seen that the system is deteriorated and the expected average time for the occurrence of a failure is 6 days ($\hat{\eta} = 6.02$) and to $\hat{\alpha} = 42.89$, it follows that the variability is $\frac{1}{42.89} = 0.023$. Then, the variability between the trucks is relatively small, but positive, giving evidence of the possible influence of factors not observed.

Table 2: Estimates for the MR model with frailty parameters via Maximum Likelihood

Parameters	MLE	SE	CI (95%)
$\hat{\beta}$	1.139	0.1012	[0.940 ; 1.337]
$\hat{\eta}$	6.026	1.6428	[2.806 ; 9.246]
$\hat{\alpha}$	42.892	6.1622	[30.815 ; 54.970]

The quality of the adjustment to the minimal repair model with frailty was evaluated by comparing the estimates of reliability function with empirical Kaplan-Meier (KM) for recurrent events. There are different estimates of the reliability function, but this work was used Wang estimator [29] it is an estimator for recurrent data correlated or independent and identically distributed times. Its function is derived from an estimator type Kaplan-Meier for recurrent events based on the intervals between times, that is $t_{i,j-1} - t_{i,j}$, called *gap times*. Consider T_{ij} the time between *j*-1-th and *j*-th of i event of the system, C_i the time " censoring times", that is, the time between the initial event and the end of the monitoring the system i and m_i content such that

$$\sum_{j=1}^{m_i-1} T_{ij} \leq C_i$$

and

$$\sum_{j=1}^{m_i} T_{ij} > C_i.$$

Thus, the term data "uncensored" refers to the set $(t_{i1}, t_{i2}, \ldots, t_{i,m_i})$ and the term "censoring" refers to the range $(t_{i1}, t_{i2}, \ldots, t^{+}_{i,m_i})$ where t^{+}_{i,m_i} refers to time between $m_i - 1$ th event and the final follow-up. So m_i is the number of recurring events for the system i.

Also consider $m_i^* = 1$ if $m_i = 1$ and $m_i^* = m_i - 1$ if $m_i \geq 2$, y_{ij} the recurrences times observed defined by $y_{ij} = t_{ij}$ if $j = 1, 2, \ldots, m_i - 1$ and $y_{ij} = t_{i,m_i}^+$ if $j = m_i$ and $S^*(t)$ the total number of systems at risk in t given by

$$S^*(t) = \sum_{i=1}^{n} \left[\frac{a_i}{m_i^*} \sum_{j=1}^{m_i^*} I(y_{ij} \geq t) \right]$$

where a_i is a function to positive values of the systems with censored values in which $E(a_i^2) < \infty$ and $I(y_{ij} \geq t)$ is 0 if $y_{ij} < t$ and 1 if $y_{ij} \geq t$.

Let

$$d^*(t) = \sum_{i=1}^{n} \left[\frac{a_i I(m_i \geq 2)}{m_i^*} \sum_{j=1}^{m_i^*} I(y_{ij} = t) \right].$$

Thus, whether $y_1^*, y_2^*, \ldots, y_K^*$ the "uncensored" times, distinct and ordered, the estimator Kaplan-Meier by Chang Wang is given

$$\hat{R}(t) = \prod_{y_i^* \geq t} \left\{ 1 - \frac{d^*(y_i^*)}{R^*(y_i^*)} \right\}.$$

This formula expresses the reliability function of the time between two successive events, known as recurrent reliability function. It is assumed that the systems have a common marginal reliability function, and sometimes within intercurrents units they are correlated.

The graph of Figure (3) shows the Kaplan-Meier estimated versus the reliability function of the minimum repair model with and without frailty. It is seen that the decay of the estimated reliability follows the Kaplan-Meier curve, suggesting that both models fit the data well.

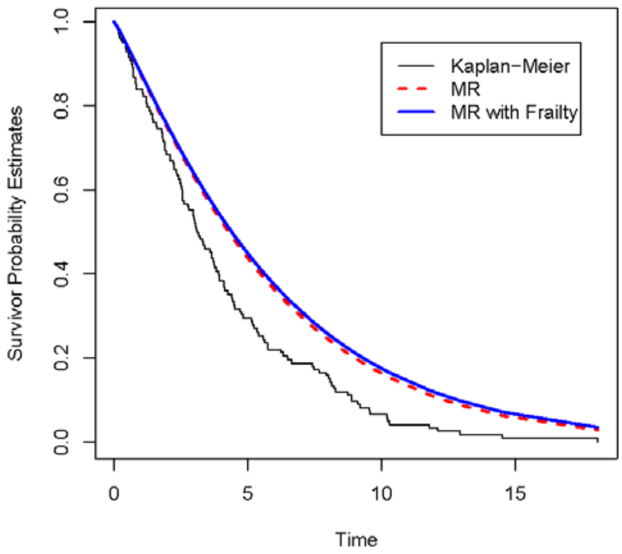


Figure 3: KM comparison chart with estimated reliability function for the model MR with frailty

It is interesting to see which model fits best, minimal repair model with or without frailty best fits the data set. For this purpose, the criterion of AIC (*Akaike Information Criterion*), BIC (*Bayesian Information Criterion*) and the log-likelihood were used. Both the AIC and for the BIC, lower values correspond to the best models and the log-likelihood function, the best model is the one with the highest value. Table (3) are the values of the AIC criteria, BIC and log-likelihood of each model. According to these criteria, the chosen model is the model of MR without frailty.

Table 3: Criteria AIC, BIC and log-likelihood of the models

Model	AIC	BIC	Log-likelihood
Without frailty	618,3623	624,0819	-307,1811
With frailty	685,8813	694,4607	-339,9406

The fact that the model MR (without frailty term) have been selected model suggests that the trucks are similar.

6. Conclusions

In this paper, a brief review of repairable systems was made and the minimal repair is shown in more detail. The main concepts of frailty were presented and the model of interest, minimal repair model with frailty, was presented and estimation method was proposed by measuring the unobserved heterogeneity between systems by inserting a latent random variable in the function failure intensity. Criteria methods were applied in order to choose the best fitted model, repairable systems with minimum repair with and without frailty. Furthermore, a simulation study was conducted for the proposed model, and the results are in agreement with that expected for maximum likelihood estimators and an application with the data of 5 dump trucks Brazilian mining company, can measure the heterogeneity observed, it was possible to show the applicability of the model.

Acknowledgment

The authors thank CAPES for financial support during the course of this project.

References

[1]. Frank Proschan. Theoretical explanation of observed decreasing failure rate. Technometrics, 5 (3): 375-383, 1963.

[2]. James W Vaupel, Kenneth G Manton, and Eric Stallard. The impact of heterogeneity in individual frailty on the dynamics of mortality. Demography, 16 (3): 439-454, 1979.

[3]. Maria Luíza Guerra de Toledo. Determination of the optimal periodic maintenance policy under imperfect repair assumption. PhD thesis, Escola de Engenharia - Departamento de Engenharia da Produçao, 2014.

[4]. Harold Ascher and Harry Feingold. Repairable systems reliability: modeling, inference, misconceptions and their causes. M. Dekker New York, 1984.

[5]. Bo Henry Lindqvist. On the statistical modeling and analysis of repairable systems. Statistical science, pages 532-551, 2006.

[6]. RI Phelps. Replacement policies under minimal repair. Journal of the Operational Research Society, 32 (7): 549-554, 1981.

[7]. Richard E Barlow and Frank Proschan. Mathematical Theory of Reliability. Society for Industrial and Applied Mathematics, 1987.

[8]. M Zhao and M Xie. On maximum likelihood estimation for a general non-homogeneous poisson process.

Scandinavian journal of statistics, pages 597-607, 1996.

[9]. Dong Ho Park, Gi Mun Jung, and Joon Keun Yum. Cost minimization for periodic maintenance policy of a system subject to slow degradation. Reliability Engineering & System Safety, 68 (2): 105-112, 2000.

[10]. Hongzhou Wang. A survey of maintenance policies of deteriorating systems. European journal of operational research, 139 (3): 469-489, 2002.

[11]. K Muralidharan. A review of repairable systems and point process models. In ProbStat Forum, volume 1, pages 26-49, 2008.

[12]. Masaaki Kijima. Some results for repairable systems with general repair. Journal of Applied probability, pages 89-102, 1989.

[13]. Odd Aalen. Nonparametric inference for a family of counting processes. The Annals of Statistics, pages 701-726, 1978.

[14]. Vera Lucia Damasceno Tomazella. Modelagem de dados de eventos recorrentes via processo de Poisson com termo de fragilidade. PhD thesis, Instituto de Ciências Matemáticas e de Computação, 2003.

[15]. David Roxbee Cox and Valerie Isham. Point processes, volume 12. CRC Press, 1980.

[16]. David Roxbee Cox and Peter AW Lewis. The statistical analysis of series of events. Monographs on Applied Probability and Statistics, London: Chapman and Hall, 1966, 1, 1996.

[17]. Larry H Crow. Reliability analysis for complex, repairable systems. in reliability and biometry, ads. f. proschan and r. j. serfling, philadelphia. Technical report, DTIC Document, 1974.

[18]. Maristela Dias De Oliveira, Enrico A Colosimo, and Gustavo L Gilardoni. Power law selection model for repairable systems. Communications in Statistics-Theory and Methods, 42 (4): 570-578, 2014.

[19]. David R Cox. Regression models and life tables (with discussion). Journal of the Royal Statistical Society, Series B, 34 (2), 1972.

[20]. Per Kragh Andersen, Ornulf Borgan, Richard D Gill, and Niels Keiding. Statistical models based on counting processes. Springer Science & Business Media, 2012.

[21]. Philip Hougaard. Frailty models for survival data. Lifetime data analysis, 1 (3): 255-273, 1995.

[22]. Debajyoti Sinha and Dipak K Dey. Semiparametric bayesian analysis of survival data. Journal of the American Statistical Association, 92 (439): 1195-1212, 1997.

[23]. David G Clayton. A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. Biometrika, 65 (1): 141-151, 1978.

[24]. David Oakes. A model for association in bivariate survival data. Journal of the Royal Statistical Society. Series B (Methodological), pages 414-422, 1982.

[25]. Philip Hougaard. Life table methods for heterogeneous populations: distributions describing the heterogeneity. Biometrika, 71 (1): 75-83, 1984.

[26]. Chris Elbers and Geert Ridder. True and spurious duration dependence: The identifiability of the proportional hazard model. The Review of Economic Studies, 49 (3): 403-409, 1982.

[27]. Jerald F Lawless. Regression methods for poisson process data. Journal of the American Statistical Association, 82(399):808–815, 1987.

[28]. Andreas Wienke. Frailty models in survival analysis. CRC Press, 2010.

[29]. Mei-Cheng Wang and Shu-Hui Chang. Nonparametric estimation of a recurrent survival function. Journal of the American Statistical Association, 94 (445): 146-153, 1999.

Published: Volume 2017, Issue 6 / June 25, 2017