Non-Parametric Tests for Imperfect Repair

GUSTAVO GILARDONI
Departamento de Estatística, UnB
Prédio CIC/EST, Campus Darcy Ribeiro
Brasília - DF, 70910-900
gilardon@gmail.com

ENRICO COLOSIMO
Departamento de Estatística, UFMG
Av. Antônio Carlos 6627 - Prédio do ICEx
Belo Horizonte - MG, 31270-901
enricoc57@gmail.com

MARIA LUÍZA TOLEDO[1]
Escola Nacional de Ciências Estatísticas, IBGE
Rua André Cavalcanti 106
Rio de Janeiro - RJ, 20231-050
malutoledo83@gmail.com

MARTA FREITAS
Departamento de Engenharia de Produção, UFMG
Av. Antônio Carlos 6627 - Escola de Engenharia
Belo Horizonte - MG, 31270-901
marta.afreitas@gmail.com

ABSTRACT

In statistical models for repairable systems, the effect of repairs after failures may be assumed as minimal or imperfect. However, no test is available to decide, based on data, which of these assumptions is true for a system. In this paper, it is proposed a general statistical test procedure in order to test the basic hypothesis *minimal* vs *imperfect repair*. Two tests are presented, which are based on the binomial and multinomial distributions. Empirical studies for the tests are presented, and shows that, under many scenarios, it has a good performance in terms of the type I and II error rates. An application with real data involving failures in trucks from a mining company is also presented.

KEYWORDS. Reliability. Repairable systems. Imperfect repair.

Paper topics: PO na Indústria. Estatística. Modelos Probabilísticos.

[1]Corresponding Author

1. Introduction

In the statistical modeling of data from repairable systems, the most explored assumptions are Minimal Repair (MR), which returns the system to the condition just before the failure, and Perfect Repair (PR), which leaves the system as if it were new. Nevertheless, a more realistic assumption for many systems is the Imperfect Repair (IR), meaning that the system returns to an intermediate state between MR and PR. Models based on such assumption have been studied by several authors, among them [Kijima et al., 1988], [Shin et al., 1996], [Pan e Rigdon, 2009], [Corset et al., 2012], and [Gilardoni et al., 2016].

In the IR context, [Toledo et al., 2015] analyzed a data set of maintenance times in a sample of five dump trucks owned by a Brazilian mining company. Figures 1 and 2 illustrate this data set.

This data set was analyzed using the ARA and ARI families of imperfect repair models [Doyen e Gaudoin, 2004] with different orders of memory. In addition, the authors proposed a goodness-of-fit plot and other procedures to select the best model and, based on it, computed reliability predictors for the trucks.

Nevertheless, in the context of model selection, it would be interesting to develop a general statistical test procedure in order to allow practitioners to answer, first of all, a very basic question: are we under a MR or an IR situation? In other words, before moving to a model selection it would be useful to test the basic hypothesis *minimal vs imperfect repair.* In this paper, non-parametric tests for this purpose are presented, based on the binomial and multinomial distributions.

The rest of the paper is organized as follows. In Section 2, the notation and some important definitions of counting processes is presented. The non-parametric tests are described in Section 3, and the results of empirical studies with these tests are presented in Section 4. Finally, Section 5 ends the paper with results for the dump trucks data set, and also with some concluding remarks.

2. Counting Processes and Notation

Assuming that failures in a repairable system are equivalently defined by the counting processes $\{N(t)\}_{t\geq 0}$, or $\{T_i\}_{i\geq 1}$, where $N(t)$ denotes the number of observed failures up to time t, T_i corresponds to the time elapsed up to the i^{th} failure, and that a repair action is taken after each failure, the distribution of such processes is completely determined by the failure intensity (or simply intensity) function defined by

$$\lambda(t) = \lim_{\delta t \to 0} \frac{P(N(t+\delta t) - N(t) = 1 | \Im_{t^-})}{\delta t}, \quad \forall t \geq 0 \tag{1}$$

where $\Im_{t^-}$ is the minimal filtration defined by the history set of all failure times occurred before t.

Under MR assumption, the counting process $\{N(t)\}_{t\geq 0}$ is a *Nonhomogeneous Poisson Process* (NHPP), and the failure intensity function (1) does not depend on the past information. This function is called the rate of occurrence of failures (ROCOF) function, given by

$$\rho(t) = \lim_{\delta t \to 0} \frac{P(N(t+\delta t) - N(t) = 1)}{\delta t}. \tag{2}$$

Consider k identical repairable systems from the same population, where failures occur independently, and assume the following conditions:

- At each failure, a repair action with negligible length time is performed.
- n_i failures are observed in the i^{th} system, $i = 1, 2, \ldots, k$.
- $N = \sum_{i=1}^{k} n_i$ is the total number of observed failures for the k systems.

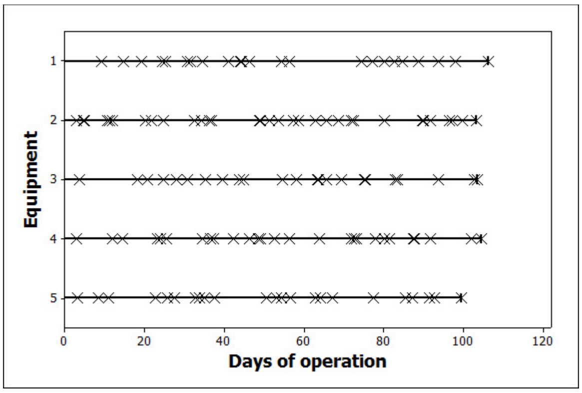

Figura 1: Failure times in days of operation for each truck (horizontal lines are trucks and “x” are failures).

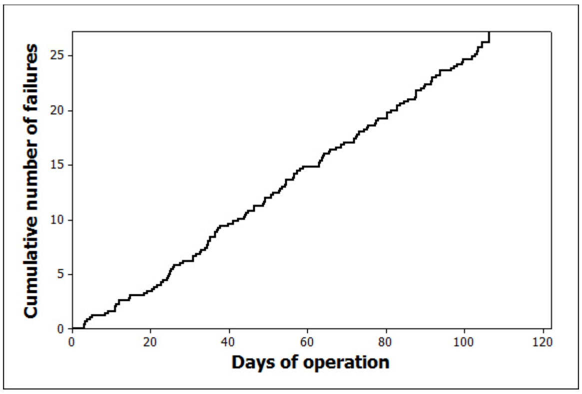

Figura 2: Cumulative number of failures *versus* days of operation in the trucks data set.

- Let $T_{i,j}$ $(i = 1, 2, \ldots, k,\ j = 1, 2, \ldots, n_i)$ be random variables representing the failure times for the i^{th} system, recorded as the time since the initial start-up of the system ($T_{i,1} < T_{i,2} < \ldots < T_{i,n_i}$). Also, let $t_{i,j}$ denote their observed values.

3. Non-Parametric Tests

The main purpose here is to test the following null and alternative hypotheses:

$$H_0\text{: MR (NHPP)} \quad \text{vs} \quad H_1 : \text{No harmful first order repair.}$$

Alternative hypothesis means the following: (1) repair improves the system and (2) just depend on the previous failure time. In an similar way, it is possible to define:

$$\lambda_1(t) = \lim_{\delta t \to 0} \frac{P(N(t+\delta t) - N(t) = 1 | t_{N(t)})}{\delta t}, \quad \forall t \geq 0 \tag{3}$$

and redefine hypotheses as

$$H_0 : \lambda_1(t) = \rho(t) \quad \text{vs} \quad H_1 : \lambda_1(t) < \rho(t) \ \text{ for some } \ t \in (0, T].$$

In this section, non parametric testes are developed for the hypotheses above. First, in Section 3.1, a naive test statistic is proposed and it will be shown that is binomial distributed under the null hypothesis. Next, in Section 3.2 is derived a multinomial extension of binomial case.

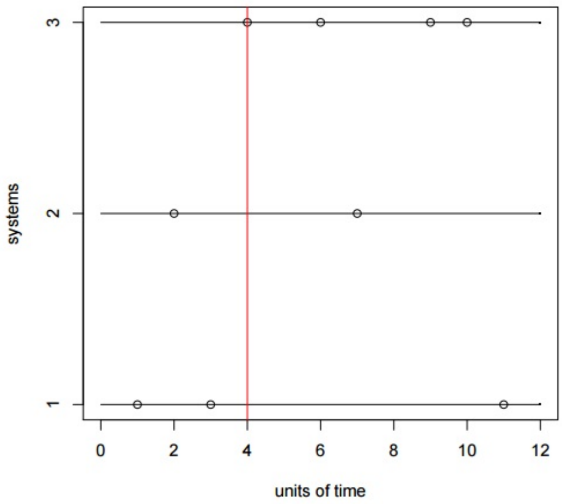

Figura 3: Failure times of three systems.

3.1. Exact Binomial Test

In order to give an intuitive idea of the test, let's consider a situation involving the follow-up of three systems. Figure 3 shows a time dot plot, where each line corresponds to a system and each "o" symbol represents a failure time.

Under the MR assumption, each system follows a NHPP and they all start at the same time as new with an increasing $\rho(t)$ in t. Consequently, due to the independent increments assumption, after the occurrence of a failure (no matter which failure) each system has the same probability of being the next one to fail.

In order to make this point clear let us observe Figure 3, in particular the vertical line at t=4 units of time. Under the null hypothesis (MR), it is fair to say that after the occurrence of this failure, the next one will have the same probability to occur in systems 1, 2 or 3.

However, under the alternative hypothesis (*no harmful repair*), if a failure occurs in a given system, the waiting time for the next failure will be longer for this system than for the other ones. In our example (Figure 3) under H_1, since the failure has occurred on system 3, it is expected that the next failure will have higher probability to occur on systems 1 or 2.

Let r_l; $l = 1, \ldots, N$, be the rank of the observed failure times $t_{i,j}$ $(i = 1, \ldots, k; j = 1, \ldots, n_i)$, in the overall sample and define the indicator vector

$$Z_{i,l} = 1\ (G_l = i)$$

for the systems membership with $G = 1, 2, \ldots, k$. In a non-parametric point of view, each observation $t_{i,j}$ is now represented by $(r_l, Z_{i,l})$.

Let's define,

$$X_l = \begin{cases} 1, \ Z_{i,l} = Z_{i,l-1}, \ \text{for all} \ \ i = 1, \ldots, k \\ 0, \ \ \text{otherwise} \end{cases}$$

for $l = 2, \ldots, N$. Note that in the illustration (Figure 3), $X_2 = X_3 = X_4 = X_6 = X_7 = X_9 = 0$ and $X_5 = X_8 = 1$.

The statistics $W = \sum_{l=2}^{N} X_l$ has a bin$(N-1, \pi)$ and under H_0, $\pi = 1/k$. Consequently, the null and alternative hypotheses are equivalent to:

$$H_0 : \ \pi = 1/k \quad \text{vs} \quad H_0 : \ \pi < 1/k.$$

and the p-value = $P(T \leq t|\pi = 1/k)$.

3.2. Multinomial Test

It is possible to generalize the binomial test. Instead of just looking at next failure it could be possible to include as many failures as possible until the process return to the same system.

Let's define the vector $X_l = (X_{l,1}, \ldots, X_{l,k})'$.

$$X_{l,1} = 1, \ \text{ if } \ Z_{i,l} = Z_{i,l-1}$$

$$X_{l,2} = 1, \qquad \text{if one diferent system } \ (Z_{i',l''} = Z_{i,l}; i \neq i') \ \text{was visited until reach for the first time to } \ Z_{i,l} = Z_{i,l'}$$

$$X_{l,k} = 1, \ \text{ if k-1 diferent systems were visited until reach for the first time to } \ Z_{i,l} = Z_{i,l'}$$

for $l' < l'' < l$ and $l = k-1, \ldots, N$. Note that in the illustration (Figure 2),

$$X_3 = (0,1,0), X_4 = (0,0,1), X_5 = (1,0,0), X_6 = (0,0,1), X_7 = (0,1,0), X_8 = (1,0,0)$$

and $X_9 = (0,0,1)$.

Let's define the vector $W = \sum_{l=k-1}^{N} X_l$. W has a multinomial $(N-k+1, \pi_1, \pi_2, \ldots, \pi_k)$ and under H_0, $\pi_1 = \pi_2 = \ldots = \pi_k = 1/k$. Consequently, the null and alternative hypotheses are equivalent to:

$$H_0 : \ \pi_1 = \pi_2 = \ldots = \pi_k = 1/k. \quad \text{vs} \quad H_1 : \pi_1 < \pi_2 < \ldots < \pi_k.$$

In the illustration in Figure 3, $W = (2,2,3)$.

4. Empirical Studies

In this section we investigate, through simulation studies, the size and power of the proposed tests.

Section 4.1 shows the results of Monte Carlo simulations implemented using as the null hypothesis a NHPP with an intensity function given by the power law intensity function. In Section 4.2, the power of these tests is evaluated under the alternative hypothesis using the ARA_1 (Arithmetic Reduction of Age of order 1) and ARI_1 (Arithmetic Reduction of Intensity of order 1) classes of models [Doyen e Gaudoin, 2004].

4.1. Descritive size

An extensively explored parametric form under the assumption of NHPP is the Power Law Process (PLP), with ROCOF function

$$\rho(t) = (\beta/\eta)(t/\eta)^{\beta-1} \tag{4}$$

and its MCF is given by $\Lambda(t) = \int_0^t \rho(u)du = (t/\eta)^{\beta}$. Here, η is a scale parameter, and β is a shape parameter. When $\beta > 1$, $\rho(t)$ increases in t, indicating that the system is deteriorating.

Small size Monte Carlo Simulations (10000 replicates) were run for some scenarios. Scenarios include $\eta = 1$, $\beta = 1.5$ and 2; T (truncation time)$= 5, 10, 15$ and k (number of systems)$= 5$ and 10. Results are presented in Tables 1 and 2, for binomial and multinomial tests, respectively.

In general, descriptive sizes are very closed to the nominal values. In particular, the descriptive sizes are sligtly smaller than the nominal ones, specially for the binomial test. This is an indication that the proposed test has a good performance in terms of the type I error rate.

Tabela 1: Monte Carlo simulation results for the empirical size of the binomial test.

Scenario			Covarage		$\widehat{\Lambda}$/System
			$\alpha = 0.05$	$\alpha = 0.010$	
$k = 5$	$T = 5$	$\beta = 1.5$	0.031	0.068	11
		$\beta = 2$	0.044	0.086	25
	$T = 10$	$\beta = 1.5$	0.039	0.085	31
		$\beta = 2$	0.046	0.090	100
	$T = 15$	$\beta = 1.5$	0.045	0.085	58
		$\beta = 2$	0.048	0.091	225
$k = 10$	$T = 5$	$\beta = 1.5$	0.035	0.074	11
		$\beta = 2$	0.039	0.083	25
	$T = 10$	$\beta = 1.5$	0.040	0.082	32
		$\beta = 2$	0.045	0.087	100
	$T = 15$	$\beta = 1.5$	0.044	0.090	58
		$\beta = 2$	0.046	0.095	225

Tabela 2: Monte Carlo simulation results for the empirical size of the multinomial test.

Scenario			Covarage		$\widehat{\Lambda}$/System
			$\alpha = 0.05$	$\alpha = 0.010$	
$k = 5$	$T = 5$	$\beta = 1.5$	0.049	0.099	10
		$\beta = 2$	0.051	0.102	24
	$T = 10$	$\beta = 1.5$	0.050	0.096	30
		$\beta = 2$	0.050	0.100	99
	$T = 15$	$\beta = 1.5$	0.050	0.099	57
		$\beta = 2$	0.052	0.098	224
$k = 10$	$T = 5$	$\beta = 1.5$	0.046	0.095	9
		$\beta = 2$	0.048	0.100	23
	$T = 10$	$\beta = 1.5$	0.046	0.094	30
		$\beta = 2$	0.046	0.097	98
	$T = 15$	$\beta = 1.5$	0.047	0.095	56
		$\beta = 2$	0.049	0.092	223

4.2. Power

ARA and ARI classes of imperfect repair models were proposed by [Doyen e Gaudoin, 2004]. ARA_1 and ARI_1 are special models of these classes of models that just depend on the last failure time. That is, ARA_1 is expressed by

$$\lambda_{ARA_1}(t) = \lambda_R(t - (1-\theta)t_{N(t)})$$

and ARI_1

$$\lambda_{ARI_1}(t) = (1-\theta)\lambda_R(t_{N(t)})$$

where $\lambda_R(t)$ is the initial/reference intensity function and θ is the repair efficiency. Power Law Process (PLP), as described in (4), is taken as λ_R.

Tables 3 to 6 show the results for binomial and multinomial tests for ARI1 using two values for $\theta = 0.1$ and 0.5. $\theta = 0.1$ means a better repair, in the sense that almost renew the system, than $\theta = 0.5$.

Results show a clear vantage of the multinomial over the binomial test. Both of them have a nice performance for $\theta = 0.1$.

5. Application and Final Remarks

The test proposed in Section 3 was applied to the data set of failures in the dump trucks, described in Section 1. This real data application gave the following results:

Tabela 3: Monte Carlo simulation results for the power at ARI ($\theta = 0.5$) of the binomial test.

Scenario			Covarage		$\widehat{\Lambda}$/System
			$\alpha = 0.05$	$\alpha = 0.010$	
$k = 5$	$T = 5$	$\beta = 1.5$	0.047	0.097	9
		$\beta = 2$	0.070	0.137	14
	$T = 10$	$\beta = 1.5$	0.055	0.116	23
		$\beta = 2$	0.070	0.135	52
	$T = 15$	$\beta = 1.5$	0.058	0.110	42
		$\beta = 2$	0.069	0.132	115
$k = 10$	$T = 5$	$\beta = 1.5$	0.051	0.106	9
		$\beta = 2$	0.077	0.147	14
	$T = 10$	$\beta = 1.5$	0.057	0.114	23
		$\beta = 2$	0.072	0.139	52
	$T = 15$	$\beta = 1.5$	0.060	0.118	42
		$\beta = 2$	0.076	0.144	115

Tabela 4: Monte Carlo simulation results for the power at ARI ($\theta = 0.1$) of the binomial test.

Scenario			Covarage		$\widehat{\Lambda}$/System
			$\alpha = 0.05$	$\alpha = 0.010$	
$k = 5$	$T = 5$	$\beta = 1.5$	0.115	0.229	6
		$\beta = 2$	0.286	0.457	7
	$T = 10$	$\beta = 1.5$	0.170	0.280	14
		$\beta = 2$	0.418	0.575	17
	$T = 15$	$\beta = 1.5$	0.188	0.308	23
		$\beta = 2$	0.413	0.572	33
$k = 10$	$T = 5$	$\beta = 1.5$	0.151	0.274	6
		$\beta = 2$	0.389	0.564	7
	$T = 10$	$\beta = 1.5$	0.209	0.341	14
		$\beta = 2$	0.495	0.653	18
	$T = 15$	$\beta = 1.5$	0.238	0.361	24
		$\beta = 2$	0.500	0.653	33

Tabela 5: Monte Carlo simulation results for the power at ARI ($\theta = 0.5$) of the multinomial test.

Scenario			Covarage		$\widehat{\Lambda}$/System
			$\alpha = 0.05$	$\alpha = 0.010$	
$k = 5$	$T = 5$	$\beta = 1.5$	0.078	0.145	8
		$\beta = 2$	0.111	0.195	13
	$T = 10$	$\beta = 1.5$	0.082	0.154	22
		$\beta = 2$	0.095	0.180	51
	$T = 15$	$\beta = 1.5$	0.075	0.149	41
		$\beta = 2$	0.086	0.159	114
$k = 10$	$T = 5$	$\beta = 1.5$	0.091	0.168	7
		$\beta = 2$	0.155	0.258	13
	$T = 10$	$\beta = 1.5$	0.090	0.174	21
		$\beta = 2$	0.131	0.228	51
	$T = 15$	$\beta = 1.5$	0.090	0.163	41
		$\beta = 2$	0.110	0.200	113

- p-value = 0.08544 for the binomial test,
- p-value = 0.07056228 for the multinomial test.

Tabela 6: Monte Carlo simulation results for the power at ARI ($\theta = 0.1$) of the multinomial test.

Scenario			Covarage		$\widehat{\Lambda}$/System
			$\alpha = 0.05$	$\alpha = 0.010$	
$k = 5$	$T = 5$	$\beta = 1.5$	0.229	0.366	5
		$\beta = 2$	0.578	0.726	6
	$T = 10$	$\beta = 1.5$	0.325	0.469	13
		$\beta = 2$	0.741	0.855	17
	$T = 15$	$\beta = 1.5$	0.344	0.494	23
		$\beta = 2$	0.755	0.861	32
$k = 10$	$T = 5$	$\beta = 1.5$	0.391	0.548	5
		$\beta = 2$	0.899	0.954	5
	$T = 10$	$\beta = 1.5$	0.562	0.703	13
		$\beta = 2$	0.976	0.991	16
	$T = 15$	$\beta = 1.5$	0.609	0.748	23
		$\beta = 2$	0.982	0.994	32

This result is an indication against the hypothesis of minimal repair for this data set. In fact, the goodness-of-fit plot proposed in [Toledo et al., 2015] also gave evidences that the PLP model with minimal repair assumption had the worst fit to these data. So, after rejecting the null hypothesis of MR, the next step would be to focus on the modeling under the IR assumption.

In the statistical analysis of data from reparable systems, a very first question that frequently arises concerns the existence of trend in the times between failures. Specific tests are often conducted to answer this question, however, they only make sense in situations where the repair is minimal. The tests proposed in this paper allow to validate this assumption in a preliminary analysis. This is very important, especially when determining preventive maintenance policies for industrial equipment.

Referências

Corset, F., Doyen, L., e Gaudoin, O. (2012). Bayesian analysis of ara imperfect repair models. *Communications in Statistics-Theory and Methods*, 41:3915–3941.

Doyen, L. e Gaudoin, O. (2004). Classes of imperfect repair models based on reduction of failure intensity or virtual age. *Reliability Engineering and System Safety*, 84:45–56.

Gilardoni, G. L., Toledo, M. L. G., Freitas, M. A., e Colosimo, E. A. (2016). Dynamics of an optimal maintenance policy for imperfect repair models. *European Journal of Operational Research*, 248: 1104–1112.

Kijima, M., Morimura, H., e Suzuki, Y. (1988). Periodical replacement problem without assuming minimal repair. *European Journal of Operational Research*, 37:194–203.

Pan, R. e Rigdon, S. (2009). Bayes inference for general repairable systems. *Journal of Quality Technology*, 41:82–93.

Shin, I., I, T. J. L., e Lie, C. H. (1996). Estimating parameters of intensity function and maintenance effect for repairable unit. *Reliability Engineering and System Safety*, 54:1–10.

Toledo, M. L. G., Freitas, M. A., Colosimo, E. A., e Gilardoni, G. L. (2015). Ara and ari imperfect repair models: Estimation, goodness-of-fit and reliability prediction. *Reliability Engineering and System Safety*, 140:107–115.