Reliability Engineering and System Safety 150 (2016) 78–88

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

journal homepage: www.elsevier.com/locate/ress

A reliability decision framework for multiple repairable units

A.H.S. Garmabaki [a,b,*], Alireza Ahmadi [a], Jan Block [a], Hoang Pham [c], Uday Kumar [a]

[a] Division of Operation and Maintenance Engineering, Luleå University of Technology, Luleå, Sweden
[b] Department of Mathematics and Computer Science, Islamic Azad University, Nour Branch, Nour 46417-38373, Iran
[c] Department of Industrial & Systems Engineering, Rutgers University, Piscataway, NJ, USA

ARTICLE INFO

Article history:
Received 24 May 2015
Received in revised form
24 January 2016
Accepted 27 January 2016
Available online 3 February 2016

Keywords:
Multiple repairable units
Reliability modelling
Homogeneity
Heterogeneity
Trend test

ABSTRACT

In practice, the analyst is often dealing with multiple repairable units, installed in different positions or functioning under different operating conditions, and maintained by different disciplines. This paper presents a decision framework to identify an appropriate reliability model for massive multiple repairable units. It splits non-homogeneous failure data into homogeneous groups and classifies them based on their failure trends using statistical tests. The framework discusses different scenarios for analysing multiple repairable units, according to trend, intensity, and dependency of the units' failure data. The proposed framework has been verified in a fleet of aircraft and in two simulated data sets. The results show a reliability model of multiple repairable units may contain a mixture of different stochastic models. Considering single reliability models for such populations may cause erroneous calculation of the time to failure of a particular unit, which can, in turn, lead to faulty conclusions and decisions. When dealing with massive and non-homogeneous multiple repairable units, the application of the proposed framework can facilitate the selection of an appropriate reliability model.

1. Introduction

When dealing with complex technical units in aviation, telecommunication and railway sectors, manufacturers and operators must ensure their fleet will meet established performance goals and quality criteria. A good reliability programme will assure the collection of important information about the system's reliability performance throughout the operation phase and direct the use of this information in the implementation of analytical and management processes [1]. Effective reliability programmes and maintenance development require proper data collection and analysis, along with the construction of reliability models to assist in the decision-making processes.

Repairable systems are those systems that can be restored to fully satisfactory performance by a method other than replacement of the entire system [2]. Reliability analysis of repairable units can be classified into parametric and non-parametric methods. Among the parametric methods, stochastic point processes, e.g. the homogeneous Poisson process (HPP), renewal process (RP), trend renewal process (TRP), branching Poisson process (BPP), and non-homogeneous Poisson process (NHPP) are used for data analysis. For more details, see Refs. [2–10].

Reliability analysis relies on historical data, and collection of these data represents the first step. Three challenges of reliability data collection are data censoring, data aggregation (pool or combine), and data with small failure events [11–13].

In some cases, the analyst is often faced with several repairable similar items which may have different reliability performances. This is due to the fact that these units may be installed in different locations and may functioning under different operating conditions, and maintained by different maintenance policies. In other words, differences in the operating environment (due to humidity, temperature, etc.) may change the pattern of failures from item to item [7]. These variations in failure patterns may lead to differences in the failure time distributions or processes of units. Hence, the population of multiple repairable units may consist of units with different failure patterns, as well as homogeneity and heterogeneity levels.

Analysis of data without considering the above factors may hinder reliability prediction. To analyse multiple repairable units in this study, we categorises the data into homogenous groups and classifies them based on their failure trends. Various trend tests proposed in the literature can be used to make groups of units based on the items' improving, deteriorating, or trend-free properties.

In fact, trend test analysis is one of the main benchmarking tools in reliability analysis. Ascher and Feingold [2] and Lindqvist [14], etc. discussed the importance of trend tests for verifying the

* Corresponding author.
E-mail address: amir.garmabaki@ltu.se (A.H.S. Garmabaki).

http://dx.doi.org/10.1016/j.ress.2016.01.020

improvement/deterioration property before using parametric models. Kvaløy and Lindqvist [19] proposed two approaches for the trend analysis of multiple repairable units. In the first, they pooled data chronologically and derived the total time on test (TTT-statistic). In the second, they introduced a combined statistical test for multiple data sets based on the combination of single unit test statistics. These approaches are applicable to trend-free categories only when homogenous data are available.

Establishing a comprehensive framework to identify candidate reliability models which are statistically valid is a challenge. This paper briefly discusses the challenges related to using the available methods for repairable units; it suggests a procedure to identify an appropriate reliability model for multiple repairable units based on a review of available trend tests, and it introduces the associated key analytical steps. It provides a procedure for grouping homogeneous units and classifies them based on their statistical trend tests in the presence of observed and unobserved heterogeneity. The proposed framework was applied in a real case involving 36 units in the aircraft fleet of an aviation company. It was also applied to two cases using simulated data.

This paper is structured as follows. Section 2 refers to common analytical trend tests to determine the existence of trends in data for single and multiple repairable units. Section 3 describes the proposed decision framework for reliability model selection. Section 4 presents numerical examples using data from a fleet of aircraft along with two examples using simulated data. Finally, Section 5 provides a conclusion.

2. Analytical trend test and underlying concepts

The study considers repairable units observed from time $t=0$, with successive failure times denoted by t_1, t_2, …. An equivalent representation of the failure process can be in terms of the counting process $\{N(t),\ t\geq 0\}$, where $N(t)$ equals the number of failures in $(0,\ t]$. It assumes that simultaneous failures are not possible. It also assume that the repair times are negligible compared to the times between failures. The study considers processes, either single or several independent similar processes, observed at (possibly different) time intervals $(0,\ T]$. It should be noted that the main possible process discussed in this paper are HPP, RP, TRP, NHPP, HNHPP and BPP.

2.1. Trend test for single repairable unit

Trend analysis is a common statistical method used to investigate the operation changes in a repairable unit over time. A trend in the pattern of failures can be monotonic or non-monotonic. In the case of a monotonic trend, the system has a concave or a convex shape. Non-monotonic trends are said to occur when trends change with time or repeat themselves in cycles. One common non-monotonic trend is the bath-tub shape trend, in which failure rate decreases at the beginning of the equipment life, tends to be constant for a period and then increases at the end [15,16].

Some trend tests are widely used in reliability studies, including the Laplace trend test, Military Handbook test, Mann test, and Anderson–Darling test; these are described in [2,5,7,17,18].

2.1.1. Laplace trend test

The Laplace trend test has a null hypothesis of "No trend" ($H0_1$) versus the alternative hypothesis of "Monotonic trend" [7,19,20], expressed as follows:

$$U=\sqrt{12.N(t_k)}\left(\frac{\sum_1^k t_i}{\tau.N(t_k)}-0.5\right)\quad \text{where}\quad \begin{cases} k=n, \tau=T & \textit{Time}-\textit{truncated} \\ k=n-1, \tau=t_n & \textit{Failure}-\textit{truncated} \end{cases} \tag{1}$$

The test statistic U has approximately a standard normal distribution, $N(0,1)$, when the $H0_1$ is true. The null hypothesis of $H0_1$ is rejected on the significant level of $\alpha\%$ if $|U|>z_{\alpha/2}$. It may be worth noting that the Laplace trend test is optimal for an NHPP with a log-linear intensity function.

2.1.2. Military Handbook Test

As in the Laplace test, the null hypothesis ($H0_2$) for the Military Handbook test is "No trend", versus the alternative "Monotonic trend," and is expressed as

$$Z=2\sum_{i=1}^{k}\ln\frac{\tau}{t_i},\text{where}\ \begin{cases} k=n, \tau=T & \textit{Time}-\textit{truncated} \\ k=n-1,\ \tau=t_n & \textit{Failure}-\textit{truncated} \end{cases} \tag{2}$$

Low-values of Z correspond to deterioration, while large values of Z correspond to improvement. The rejection criteria of null hypothesis ($H0_2$) are given by:

$$\text{Rejection}\quad H0_2:\begin{cases} Z<\chi^2_{2n,1-\alpha/2}\ \ or\ \ Z>\chi^2_{2n,\alpha/2} & \textit{Time}-\textit{truncated} \\ Z<\chi^2_{2(n-1),1-\alpha/2}\ \ or\ \ Z>\chi^2_{2(n-1),\alpha/2} & \textit{Failure}-\textit{truncated} \end{cases} \tag{3}$$

Note that the Military Handbook test is optimal for the power law intensity function [7,21,22].

2.1.3. The Mann Test

The null hypothesis ($H0_3$) for this non-parametric test is an RP versus a monotonic trend. This trend test is calculated by counting the reverse arrangements among the times between failures as:

$$M=\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}I(X_i<X_j) \tag{4}$$

where $X_1, X_2, \ldots, X_n$ is inter-arrival times of n failures and $I(X_i<X_j)=1$ if the event $X_i<X_j$ occurs and 0 otherwise. The test statistic of Mann test is given by:

$$M_{Mann}=\frac{M-n(n-1)/4}{\sqrt{(2n^3+3n^2-5n)/72}} \tag{5}$$

The Mann test statistic is approximately distributed as a standard normal distribution; for more details, see [23]. The null hypothesis ($H0_3$) is rejected on the level of $\alpha\%$ if $|M_{Mann}|>z_{\alpha/2}$.

2.1.4. Anderson–Darling test

The Anderson–Darling test (AD) rejects the null hypothesis ($H0_4$ is "No trend") in the presence of both monotonic and non-monotonic trends when the value of AD is large.

$$AD=-\frac{1}{n}\left[\sum_{i=1}^{n}(2i-1)\left(\ln\left(\frac{T_i-T_0}{T-T_0}\right)+\ln\left(1-\frac{T_{n+1-i}-T_0}{T-T_0}\right)\right)\right]-n \tag{6}$$

The null hypothesis ($H0_4$) is rejected at the level of 5% if $|AD>|2.492$; for more information, see [6,19,24]. Besides the AD test other tests, such as the generalized AD, $V1$, $V2$, $V3$ and $V4$ tests [25], can be used to identify non-monotonic trends. The AD test and $V1$, $V2$ and $V3$ tests have HPP, while the generalized AD test and $V4$ have RP as the null hypothesis of test statistics. Interested readers are referred to [16,25,26]. Moreover, in some case, combinations of the tests are required to identify the existence of a trend. For example, if the Military Handbook and the Laplace tests reject the null hypothesis, the data do not follow an HPP. However, the data can still be trend-free [14]. These issues are discussed in Section 3.

Significant level is one of the important issues in the testing of the null hypothesis. Selection of significant level can be affected by sample size and expected losses. Hypothesis testing without considering the potential losses is not ethically and economically defensible [27]. Leamer [28] shows how the optimal significance