Reliability Engineering and System Safety 115 (2013) 111–123

Contents lists available at SciVerse ScienceDirect

Reliability Engineering and System Safety

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

Choosing an optimal model for failure data analysis by graphical approach

Tieling Zhang *, Richard Dwight

School of Mechanical, Materials and Mechatronic Engineering, University of Wollongong, NSW 2522, Australia

A R T I C L E I N F O

Article history:
Received 31 July 2012
Received in revised form
13 January 2013
Accepted 7 February 2013
Available online 7 March 2013

Keywords:
Weibull models
Failure data analysis
Model selection
Parameter estimation
Weibull probability paper

A B S T R A C T

Many models involving combination of multiple Weibull distributions, modification of Weibull distribution or extension of its modified ones, etc. have been developed to model a given set of failure data. The application of these models to modeling a given data set can be based on plotting the data on Weibull probability paper (WPP). Of them, two or more models are appropriate to model one typical shape of the fitting plot, whereas a specific model may be fit for analyzing different shapes of the plots. Hence, a problem arises, that is how to choose an optimal model for a given data set and how to model the data. The motivation of this paper is to address this issue.

This paper summarizes the characteristics of Weibull-related models with more than three parameters including sectional models involving two or three Weibull distributions, competing risk model and mixed Weibull model. The models as discussed in this present paper are appropriate to model the data of which the shapes of plots on WPP can be concave, convex, S-shaped or inversely S-shaped. Then, the method for model selection is proposed, which is based on the shapes of the fitting plots. The main procedure for parameter estimation of the models is described accordingly. In addition, the range of data plots on WPP is clearly highlighted from the practical point of view. To note this is important as mathematical analysis of a model with neglecting the applicable range of the model plot will incur discrepancy or big errors in model selection and parameter estimates.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

Modeling a given set of data by the graphical approach is an intuitive and fast way to formulate the data. The graphical approach is based on plotting data on probability papers such that normal distribution probability paper, log-normal probability paper, Weibull probability paper (WPP), etc. have been developed and widely applied. Among them, the WPP is more frequently utilized in data analysis as a Weibull distribution is appropriate to model failure times and it is flexible in modeling as such the corresponding failure rate can be decreasing, increasing, constant or other forms. If a set of data plotted on WPP is roughly scattering on a straight line, one can model the data as coming from the two-parameter Weibull distribution. If not, one can try three-parameter Weibull models or models involving multiple Weibull distributions or other types of distributions instead.

A large number of Weibull-related models have been developed, which are applied to modeling the data whose fitting plots on WPP take different shapes. These include modified Weibull distribution and its extension models [1–5], exponentiated Weibull family [6–8], mixture models [9–10], competing risk models, multiplicative models and sectional models [11–14]. These models can be utilized to analyze a given data set whose fitting plot on WPP is concave, convex, S-shaped or further other shapes. An overview on the Weibull models can be found in [15]. In recent years, there are many research papers published on the extended Weibull and modified Weibull distributions and their applications, see for example, [5,16–22]. The interest is that each of the distribution models can present a hazard function that is decreasing, increasing or bathtub shaped. While the model plot on WPP of the modified Weibull given by Lai et al. [1] and the modified Weibull extension proposed by Xie et al. [2] shows a concave curve, the extended Weibull distribution given by Marshall and Olkin [3] presents a model plot that is S-shaped or inversely S-shaped [5]. A further Weibull extension model e.g., [18] or a generalized modified Weibull distribution e.g., [19,22] with four parameters can provide more versatile properties in terms of probability density function and hazard function for model application. However, the characteristics of the model plot on WPP of these newly developed models have not been discussed. Murthy et al. [23] present the method for Weibull-related model selection with a list of commonly used distribution models but there is not a discussion in detail on the characteristics of the model plots. They first categorize the shapes of pdf, hazard function and WPP plots and then identify the category which each model belongs to.

* Corresponding author. Tel.: +6124 221 4821.
E-mail addresses: zhangtling@yahoo.com, tieling@uow.edu.au (T. Zhang), radwight@uow.edu.au (R. Dwight).

0951-8320/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.ress.2013.02.004

Notation

$f(t)$, $F(t)$	[pdf, Cdf] for a distribution that may involve sub-populations
i	index to sub-population i, $i=0, 1, 2$ unless otherwise specified
$R(t)$	reliability function (survivor function (Sf))
$h(t)$, $h_i(t)$	hrf of a distribution and its *ith* sub-population
$f_i(t)$, $F_i(t)$, $R_i(t)$	[pdf, Cdf, Sf] of sub-population i
C_f	fitting plot: $y(x)$ *vs* x
W_j	the *jth* section of C_f, $j=1, 2, 3$
η_i, β_i	[scale, shape] parameter of $R_i(t)$, all are positive
γ	location parameter of three-parameter Weibull distribution
x	$\ln(t)$
$y(t)$	$\ln(\ln-(R(t)))$
$y(x)$	$\ln(\ln-(R(e^x)))$
L_i	a straight line, $y_i(x)=\beta_i(x-\ln(\eta_i))$
I	intersection of L_1 & L_2
II	intersection of L_0 & L_1 or L_0 & L_2
L_a	asymptote to C_f as $x \to -\infty$
y', y''	[first, second] derivative of $y(x)$
$(x_{\mathrm{I}}, y_{\mathrm{I}})$	coordinates of point I in the x–y plane
$(x_{\mathrm{II}}, y_{\mathrm{II}})$	coordinates of point II in the x–y plane

The shapes of each model plots are categorized based on the mathematical background of the model. The readers cannot find the characteristics of the models in detail and then they have to read the original papers that present the models in order to have a good understanding of the models and apply them to data modeling. Lai et al. [24] give a review paper that reviews the properties of the basic Weibull distribution and lists the various extensions of the Weibull distribution. It describes the use of Weibull probability plots as a tool for model selection and briefly discusses the parameter estimation and model validation. However, this is a short overview paper and there is not a discussion in detail on the characteristics of the WPP plots of each model. In addition, it does not describe the way to select an optimal model for modeling a given data set.

It is of interest to discuss model selection and associated parameter estimation based on the data plot on WPP. Most models are appropriate for modeling the data whose fitting plot on WPP shows a concave or convex curve. These models include 3-parameter Weibull models and models involving two or three Weibull distributions.

The models with multiple Weibull distributions are more flexible in application to modeling the given data set of which the fitting plot on WPP can take S-shape or further other shapes. The models as discussed in this paper include competing risk model, multiplicative model, mixed Weibull and sectional model. Each of them is reviewed shortly.

The competing risk model was discussed in [11,25]. A physical justification for the model is that an item failure occurs due to more than one cause or failure mode, and these causes and failure modes are statistically independent. The item fails whenever a failure mode occurs. The competing risk analysis has many applications, see, e.g., [26–28] and a thorough review was given in [29].

As for multiplicative model, an interpretation to such a model is that a system consists of n components connected in parallel and such that the time to failure of the system depends on the maximum of $\{T_1, T_2, \ldots, T_n\}$ where T_i is distributed according to $F_i(t)$. Here, $F_i(t)$ is cumulative distribution function of component i. As a result, we have the model $F(t)=F_1(t)\times F_2(t)\times \ldots F_n(t)$. If such a model involves two Weibull distributions of each with two parameters, the model plot on WPP is a convex curve, see, e.g., [11].

Mixed distribution model such as the mixed Weibull distribution has been applied in industry for many years. Essentially, a mixed distribution is a distribution comprised of a number of distinct sub-distributions that have been "patched together" to form one continuous function [30]. The mixed distribution is useful when modeling the data set that can be divided into subgroups and the data in one subgroup can be treated as coming from one subpopulation because of the same failure cause. It has been recognized for more than four decades that the mixed Weibull distribution is an appropriate distribution to use in modeling the lifetimes of units that have more than one failure cause [31]. Jiang and Kececioglu [9,32], Jiang and Murthy [10], Kececioglu and Wang [31] and Ling et al. [33] studied the parameter estimation of the model by maximum likelihood estimate, the method of Least Squares and the graphical approach. The model plot on WPP of the mixed Weibull involving two distributions shows a complex pattern that has two inflection points.

The more flexible models which are appropriate to analyze complex data are sectional ones. In a sectional model (also called composite model, piece-wise model, or step function model), the failure distributions over different time intervals are given by different distribution functions. The main possible reasons to use sectional models are as follows [12]:

1. It is mathematically tractable and yields a bathtub shape for the failure rate function.
2. Its flexibility allows modeling complex data set.
3. When the material properties of an item change significantly after a certain length of time in application, then failures of the item before and after the change should be modeled by different distributions. Thus, in the case, a sectional model is appropriate to model the combined failures.

A literature review on sectional models can be referred to [12,13]. Mann, et al. [25] and Elandt-Johnson and Johnson [34] discussed the sectional model involving two Weibull distributions. Jiang and Murthy [12,14] studied the sectional models involving two Weibull distributions and parametric properties of these models. Furthermore, other forms of the sectional models involving three Weibull distributions were proposed by Jiang and Murthy [13], Zhang and Ren [35]. The shapes of plots on WPP of the sectional models with three Weibull distributions are very flexible such as those with S-shaped, three sections, concave and convex curves [35]. Sectional models have been and will be applied to many applications.

As summarized above, one typical shape of data plots can be modeled by different models and on the other hand one distribution model may be used to model different shapes of the data plots. From this, it is of interest to us that how to choose an optimal model to model a given data set based on plotting the data on WPP.

The purpose of this paper is to summarize the characterization of plots on WPP defined by different models and to present a basic procedure for choosing an optimal model to formulate the given set of data and the method for parameter estimates. The reasonability for model application is discussed from the practical point of view, as some analysts may depend on or overly emphasize the model characterization in mathematics but neglect its applicability

Download English Version:

https://daneshyari.com/en/article/7195908

Download Persian Version:

https://daneshyari.com/article/7195908

Daneshyari.com