2013

Choosing an optimal model for failure data analysis by graphical approach

Tieling Zhang
University of Wollongong, tieling@uow.edu.au

Richard Dwight
University of Wollongong, radwight@uow.edu.au

Publication Details

Zhang, T. & Dwight, R. (2013). Choosing an optimal model for failure data analysis by graphical approach. Reliability Engineering and System Safety, 115 111-123.

Choosing an optimal model for failure data analysis by graphical approach

Abstract

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.

Keywords

model, optimal, data, approach, analysis, failure, choosing, graphical

Disciplines

Engineering | Science and Technology Studies

Publication Details

Zhang, T. & Dwight, R. (2013). Choosing an optimal model for failure data analysis by graphical approach. Reliability Engineering and System Safety, 115 111-123.

Tieling Zhang and Richard Dwight. Choosing an optimal model for failure data analysis by graphical approach, ***Reliability Engineering and System Safety***, Vol. 115 (2013), pp. 111–123.

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
Email: tieling@uow.edu.au, radwight@uow.edu.au

ABSTRACT

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.

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

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	$\ell n\,(t)$
$y(t)$	$\ell n\,(-\ell n\,(R(t)))$
$y(x)$	$\ell n\,(-\ell n\,(R(e^x)))$
L_i	a straight line, $y_i(x) = \beta_i(x-\ell n\,(\eta_i))$
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_I, y_I)	coordinates of point I in the x-y plane
(x_{II}, y_{II})	coordinates of point II in the x-y plane

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. 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 modelling. 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 in practice so as to incur errors or discrepancy in model selection and parameter estimation.

The outline of this paper is as follows: Section 2 discusses the general range of WPP plots from the point view of real application; Section 3 summarizes the properties of plots defined by different models that involve two or three Weibull distributions; Section 4 discusses the method for choosing an optimal model for modeling a given set of data, which is based on the data plot on WPP; Section 5 shows two examples and Section 6 gives remarks on application of the graphical approach.

2. WPP plot and its property

The two- and three-parameter Weibull distributions are given in (1) and (2):

$$F(t) = 1 - R(t) = 1 - exp[-(t/\eta)^{\beta}], \quad (1)$$

$$F(t) = 1 - R(t) = 1 - exp\{-[(t-\gamma)/\eta]^{\beta}\}. \quad (2)$$

Using the following transformations

$$x = \ell n\,(t) \quad \text{and} \quad y = \ell n\,(-\,\ell n\,(R(t))), \tag{3}$$

one can plot *y vs x* on WPP. Equation (3) is called the Weibull transformation. By using equation (3) in equation (1), the following is obtained,

$$y = \beta\,(x - \ell n\,(\eta)). \tag{4}$$

This is a straight line. The slope of this line is β and the intercept on the x-axis is $\ell n\,(\eta)$. Similarly, by applying equation (3), equation (2) is transformed into the following:

$$y = \beta\,[\,\ell n\,(t - \gamma) - \ell n\,(\eta)] = \beta\,[\,\ell n\,(\,e^{x} - \gamma) - \ell n\,(\eta)]. \tag{5}$$

As the above, plotting *y vs x* on WPP yields a convex curve as shown in Fig. 1. This curve has the following characteristics:

$$dy\,/\,dx = \beta\,[1 + \gamma/(e^{x} - \gamma)], \tag{6}$$

$$\lim_{x \to \ell n(\gamma)} dy\,/\,dx = \infty\,, \tag{7}$$

$$\lim_{x \to \infty} dy\,/\,dx = \beta, \tag{8}$$

and

$$d^2y/dx^2 = -\beta\gamma\,e^{x}/(e^{x} - \gamma)^2 < 0. \tag{9}$$

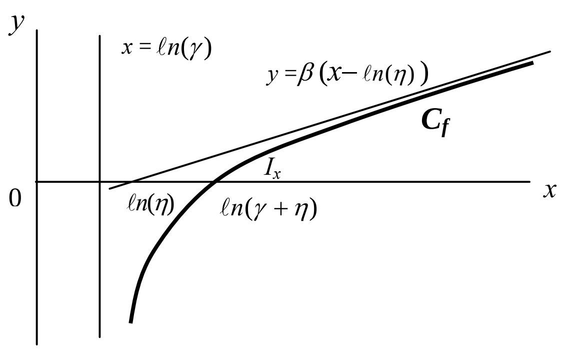


Fig. 1. Plot of 3-parameter Weibull distribution on WPP

There are two asymptotes:

$$y = \beta\,[x - \ell n(\eta)] \qquad \text{as} \qquad x \to \infty$$

and

$$x = \ell n(\gamma) \qquad \text{as} \qquad x \to \ell n(\gamma).$$

Note that C_f intersects the x-axis at $x = \ell n(\gamma + \eta)$. Let I_x denote this point, the slope of C_f at this point is

$$dy/dx\Big|_{x = \ell n(\gamma + \eta)} = \beta(1+\gamma/\eta). \quad (10)$$

For a given set of n data, the plotting procedure and parameter estimates can be performed in the following steps.

#1. Rearrange the data in increasing order, like t_1 , t_2, …, t_i, …, t_n.

#2. Compute x_i and y_i, $1 \le i \le n$, as follows:

$$x_i = \ell n\ (t_i) \quad \text{and} \quad y_i = \ell n\ \{-\ell n\ [1 - i\ /\ (n+1)]\} \quad (11)$$

or

$$x_i = \ell n\ (t_i) \quad \text{and} \quad y_i = \ell n\ \{-\ell n\ [1 - (i - 0.3)\ /\ (n+0.4)]\}$$

if n is smaller.

#3. Plot y_i *vs*. x_i on WPP.

If the data points scatter roughly along a line, this set of data can be adequately modeled by a two-parameter Weibull distribution. If not, one can try a three-parameter Weibull distribution or, otherwise, other models as will be discussed in the following section.

If the data set (t_1 , t_2, …, t_i, …, t_n) contains censored data, #2 ought to be modified. Refer to [36] for description in detail on the modifications.

Plotting a given set of data on WPP is based on the Weibull transformation. From equation (11), we know that $y_i \in [\ell n\ (-\ell n(1 - 1/(n+1))),\ \ell n\ (-\ell n(1 - n/(n+1)))]$ where n is the total number of data in a set. When n = 1000, for example, $y_i \in [-6.90826, 1.93279]$, that means, the fitting plot is in the range of $y \in [-6.90826, 1.93279]$. Similarly, if n = 100, $y \in [-4.61015, 1.52934]$. Table 1 gives the range of the plot versus the data sample size, n. Because the minimum and maximum values of y depend on n, on the other hand, if y = 3.0, then n is larger than 5.28×10^8. How large the sample size is!

To note the range of fitting plot on WPP in general is important. As without considering this will incur big error or discrepancy in model selection and parameter estimates. Unfortunately, this was neglected in some research papers when discussing the model properties and associated parameter estimation by the graphical approach.

Table 1 Values of y_1 and y_n ($y_n = \ell n\{-\ell n\ [1 - n\ /\ (n+1)]\}$)

n	10^6	10^5	10^4	10^3	10^2	50	20
$F(t_n)$	9.99999×10^{-7}	9.9999×10^{-6}	9.999×10^{-5}	9.99001×10^{-4}	9.90099×10^{-3}	0.0196078	0.047619
y_n	2.62579	2.44347	2.22034	1.93279	1.52934	1.3691	1.11334

$F(t_1)$	0.999999	0.99999	0.9999	0.999001	0.990099	0.980392	0.952381
y_1	-13.8155	-11.5129	-9.21039	-6.90826	-4.61015	-3.92194	-3.02023

Based on these discussions, we examine different shapes of plots on WPP and associated models for fitting the plots. Parameter estimates of the models are also presented.

3. Plots on WPP and associated models

In this section, we discuss different shapes of WPP plots and present the corresponding models that are appropriate to fit. These plot shapes are concave, convex, S-shaped or inversely S-shaped. Accordingly, the models are of competing risk model, multiplicative model, mixed Weibull model and sectional model. The main points on parameter estimates of these models are also presented.

3.1. Convex curve and associated models

3.1.1. Sectional models involving two Weibull distributions

Look at Fig. 2, the left side of the plot is quite like on a straight line. On the other hand, when x tends to be very large, the right side of the plot tends to another straight line. The sectional models involving two Weibull distributions are appropriate to model this plot. Here, we describe the relevant models.

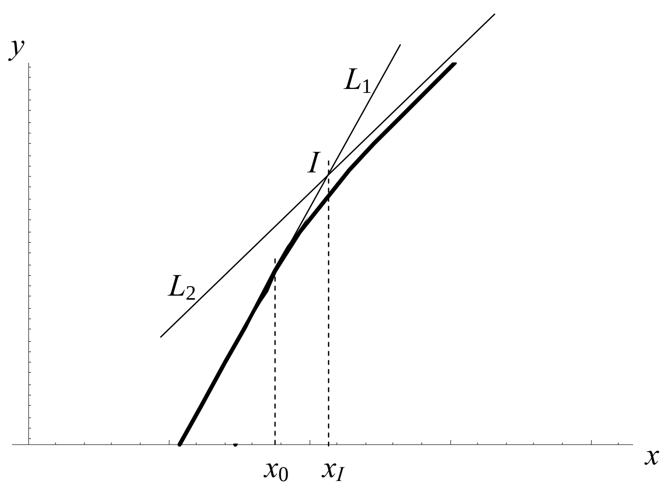


Fig. 2. A typical plot for sectional models ($\beta_1 > \beta_2$)

Model-*a*:

The Cdf is

$$F(t) = \begin{cases} 1 - exp[-(t/\eta_1)^{\beta_1}], & 0 \le t \le t_0; \\ 1 - exp\{-[(t-\gamma)/\eta_2]^{\beta_2}\}, & t_0 < t < \infty. \end{cases} \tag{12}$$

Its pdf and reliability function are imposed continuity at $t = t_0$, the parameters are constrained to satisfy the following relations:

$$t_0 = [\eta_1^{\beta_1}(\beta_2/\beta_1/\eta_2)^{\beta_2}]^{1/(\beta_1-\beta_2)}, \tag{13}$$

$$\gamma = (1-\beta_2/\beta_1)\, t_0. \tag{14}$$

It needs $\beta_1 > \beta_2$ for the model. Using the Weibull transformation, from equation (12), we have

$$y = y(x) = \begin{cases} \beta_1[x - \ell n(\eta_1)] & -\infty < x \le \ell n(t_0), \\ \beta_2[\ell n(e^x - \gamma) - \ell n(\eta_2)] & \ell n(t_0) < x < \infty. \end{cases} \tag{15}$$

As a result, C_f is a straight line for $-\infty < x \le \ell n(t_0) = x_0$, which is identical with L_1, and a smooth curve for $x_0 < x < \infty$ as shown in Fig. 2. As $x \to \infty$, the asymptotic slope of C_f is given by β_2 and hence the asymptote is L_2,

$$y(x) = \beta_2[x - \ell n(\eta_2)]. \tag{16}$$

Model-*b*:

The Cdf is given by

$$F(t) = \begin{cases} F_1(t) & \text{for } 0 \le t \le t_0, \\ 1 - k\, R_2(t) & \text{for } t_0 < t < \infty \end{cases} \tag{17}$$

where, $F_1(t) = 1 - exp[-(t/\eta_1)^{\beta_1}]$ and $R_2(t) = exp[-(t/\eta_2)^{\beta_2}]$, k is a parameter with $k > 0$. Or, this model is given in another form

$$R(t) = \begin{cases} R_1(t), & 0 \le t \le t_0; \\ k\, R_2(t), & t_0 < t < \infty. \end{cases}$$

Impose the continuity hold for the pdf and reliability function at $t = t_0$, the parameters are constrained to satisfy the following two equations:

$$t_0 = [\beta_1\eta_2^{\beta_2}/\beta_2/\eta_1^{\beta_1}]^{1/(\beta_2-\beta_1)}, \tag{18}$$

$$k = exp[(1-\beta_2/\beta_1)(t_0/\eta_2)^{\beta_2}]. \tag{19}$$

As a result, the model has 4 independent parameters other than 6. This model requires $\beta_1 \neq \beta_2$, otherwise, it reduces to a single Weibull distribution. From equation (19), $k > 1$ if $\beta_1 > \beta_2$ and $k < 1$ if $\beta_1 < \beta_2$.

The model plot on WPP is as follows:

$$y = y(x) = \begin{cases} \beta_1 [x - \ell n(\eta_1)], & -\infty < x \leq \ell n(t_0) \\ \beta_2 [x - \ell n(\eta_2)] + \ell n[1 - \eta_2^{\beta_2} \ell n(k) e^{-\beta_2 x}], & \ell n(t_0) < x < \infty. \end{cases} \tag{20}$$

This defines that C_f is a straight line for $-\infty < x \leq \ell n(t_0) = x_0$, which is identical with L_1, and a smooth curve for $x_0 < x < \infty$ as indicated in Fig. 2 with $\beta_1 > \beta_2$. If $\beta_1 < \beta_2$ ($k < 1$), C_f is concave for $x_0 < x < \infty$. As $x \to \infty$, the asymptotic slope of C_f is given by β_2, as a result, this asymptote is L_2 given by (16).

The differences for model plots between Model-*a* and Model-*b* are

$$x_0 = x_I - [\beta_2 \ell n(\beta_1 / \beta_2)] / (\beta_1 - \beta_2), \tag{21}$$

and

$$y_0 = y(x)\Big|_{x=x_0} = y_I - [\beta_1 \beta_2 \ell n(\beta_1 / \beta_2)] / (\beta_1 - \beta_2) \tag{22}$$

for Model-*a*, while

$$x_0 = x_I - [\ell n(\beta_1 / \beta_2)] / (\beta_1 - \beta_2) \tag{23}$$

and

$$y_0 = y(x)\Big|_{x=x_0} = y_I - [\beta_1 \ell n(\beta_1 / \beta_2)] / (\beta_1 - \beta_2) \tag{24}$$

for Model-*b*. These conditions are used to judge which model is better to model the data plot as shown in Fig. 2. The discussion in detail about Model-*a* and Model-*b* can be referred to [11].

Model-*c*

This model is given by

$$F(t) = \begin{cases} kF_1(t), & 0 \leq t \leq t_0; \\ F_2(t), & t_0 \leq t. \end{cases} \tag{25}$$

As $F(t)$ and $f(t)$ are required to be continuous at $t = t_0$, there are the following two equations:

$$kF_1(t_0) = F_2(t_0) \quad \text{and} \quad kf_1(t_0) = f_2(t_0). \tag{26}$$

By defining

$$z_1 = (t/\eta_1)^{\beta_1},\ z_2 = (t/\eta_2)^{\beta_2},\ c = (\eta_1/\eta_2)^{\beta_2} \text{ and } \beta = \beta_2/\beta_1,$$

equation (26) is changed into the following two equivalent ones.

$$exp(c z^{\beta}) - c\beta z^{\beta-1}(e^z - 1) - 1 = 0, \tag{27}$$

and

$$k = \frac{1-\exp(-cz_0^{\beta})}{1-\exp(-z_0)} \tag{28}$$

where $z \equiv z_1 = (t/\eta_1)^{\beta_1}$, and $z_0 = z_1(t_0) = (t_0/\eta_1)^{\beta_1}$ is the nonzero solution of equation (27). The characteristics of the model are as follows.

i. This model requires $\beta_1 \neq \beta_2$, otherwise, it reduces to a single Weibull distribution.

ii. When $\beta_1 < \beta_2\, (\beta > 1)$, equation (27) has a nonzero solution z_0 ($z_0 > 0$). As $z_0 = (t_0/\eta_1)^{\beta_1}$, then $t_0 = \eta_1\, z_0^{1/\beta_1}$. It is proved that $t_0 = \eta_1\, z_0^{1/\beta_1} > \eta_1\, c^{1/(\beta_1-\beta_2)}$ and $k > 1$ under $\beta > 1$. Details for the proof can be referred to [12].

iii. When $\beta_1 > \beta_2\, (\beta < 1)$, equation (27) has a nonzero solution z_0 ($z_0 = (t_0/\eta_1)^{\beta_1}$), hence, $t_0 > \eta_1\, c^{1/(\beta_1-\beta_2)}$ and $k < 1$. Refer to [12] for details of the proof.

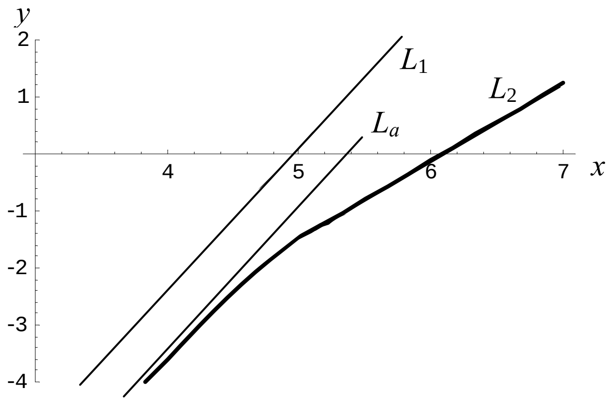


Fig. 3. WPP plot of Model-*c* for $\beta_1 > \beta_2$ ($\beta < 1$)
$\beta_1 = 2.5$, $\eta_1 = 141$; $\beta_2 = 1.35$, $\eta_2 = 440$; $t_0 = 158.606$, $k = 0.3018$.

By using the Weibull transformation in equation (3), from equation (25) we have the following:

$$y = y(t) = \begin{cases} \ell n\,(-\ell n\,(1-k+k\,exp(-(e^x/\eta_1)^{\beta_1})), & -\infty < x \le x_0 \\ \beta_2(x - \ell n\,(\eta_2)), & x_0 < x < \infty. \end{cases} \tag{29}$$

As known that $\ell n\,(1-u) \approx -u$ for small u, then

$$y \approx \ell n\,(k) + \beta_1(x - \ell n\,(\eta_1)) = \beta_1 x + \ell n\,(k/\eta_1^{\beta_1}) \tag{30}$$

as $x \to -\infty$ (or $t \to 0$). This implies that the WPP plot has an asymptote L_a given by (30), which is parallel to L_1 but displayed vertically by $|\ell n(k)|$ as $x \to -\infty$. L_a locates above L_1 when $\beta > 1$ and below L_1 when $\beta < 1$. See Figs. 3 and 10 (given in Section 3.2.1.2) for the two cases. The WPP plot is convex when $\beta < 1$ and concave when $\beta > 1$, and it is identical with L_2 in the range of $x > x_0$ ($x_0 = \ell n\,(t_0)$).

3.1.2. *Multiplicative model*

In general, an n-fold multiplicative model is given by

$$F(t) = F_1(t) \times F_2(t) \times \ldots \times F_n(t) \tag{31}$$

where, $F(t)$ and $F_i(t)$ (i = 1 to n) are cumulative distribution functions, respectively.

Here, we focus on the 2-fold multiplicative model which is described by

$$F(t) = F_1(t) \times F_2(t) \tag{32}$$

where $F_i(\mathrm{t}) = 1- exp[-(t/\eta_i)^{\beta_i}]$, $t \geq 0$, i = 1, 2. It is named the 2-fold Weibull multiplicative model. Without loss of generality, assume that $\beta_1 \leq \beta_2$ and $\eta_1 > \eta_2$ for the case $\beta_1 = \beta_2$.

From equation (32), the reliability function is

$$R(t) = R_1(t) + R_2(t) - R_1(t)R_2(t)\,. \tag{33}$$

Using the Weibull transformation (3), we have

$$y = \ell n\,\{-\ell n\,[R_1(e^x) + R_2(e^x) - R_1(e^x)R_2(e^x)]\}. \tag{34}$$

Hence, equation (34) gives a smooth curve C_f. As the derivation given in [11], there are the followings

$$\ell im_{x \to -\infty}\, dy(x)/dx \;=\; \beta_1 + \beta_2$$

and

$$\ell im_{x \to \infty}\, dy(x)/dx \;=\; \beta_1\,.$$

This implies C_f is a convex curve, see for example, Figs. 4 ~ 6. Note that C_f has two asymptotes. One is straight line L_a given by

$$y(x) = \beta_1\,[x - \ell n(\eta_1)\,] + \beta_2\,[x - \ell n(\eta_2)\,] \tag{35}$$

as $x \to -\infty$ and the other is line L_1 as $x \to \infty$. Equation (35) shows that there exists an interesting relation between L_a, and L_1 and L_2.

Let $A(x_A\,, y_A)$ and $B(x_B, y_B)$ denote the intersections of L_a with L_1 and L_2, respectively. Hence,

$$x_A = \ell n(\eta_2)\,,\quad y_A = \beta_1\,\ell n(\eta_2/\eta_1) \tag{36}$$

$$x_B = \ell n(\eta_1)\,,\quad y_B = \beta_2\,\ell n(\eta_1/\eta_2)\,. \tag{37}$$

By noting that I is the intersection of L_1 and L_2, the relationship among A, B and I can be derived, that depends on the parameter values of the distribution functions involved. First, to consider the case of $\beta_1 \neq \beta_2$ and then the case of $\beta_1 = \beta_2$.

3.1.2.1. Case of $\beta_1 \neq \beta_2$

Case (*i*): $\eta_1 << \eta_2$ (well separated case)

In this case, there exists

$$dy(x) / dx\Big|_{x=\ell n(\eta_2)} \cong \beta_2 . \tag{38}$$

This implies that the tangent to C_f at $x = \ell n(\eta_2)$ is approximately overlapping with line L_2. A typical plot of y *vs.* x is shown in Fig. 4. This is interesting that $x_B < x_A < x_I$.

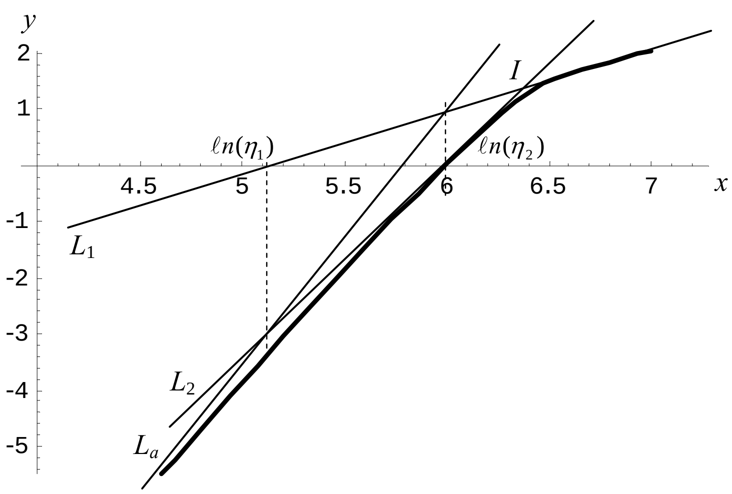


Fig. 4. $\beta_1 = 1.05$, $\eta_1 = 155$; $\beta_2 = 3.5$, $\eta_2 = 389$ $(\eta_1 << \eta_2)$

Case (*ii*): $\eta_1 >> \eta_2$ (well separated case)

In this case, we have

$$dy(x) / dx\Big|_{x=\ell n(\eta_1)} \cong \beta_1 . \tag{39}$$

The implication of this is that the tangent to C_f at $x = \ell n(\eta_1)$ is approximately overlapping with Line L_1. A typical plot of y *vs.* x is shown in Fig. 5. The important to note is that $x_I < x_A < x_B$.

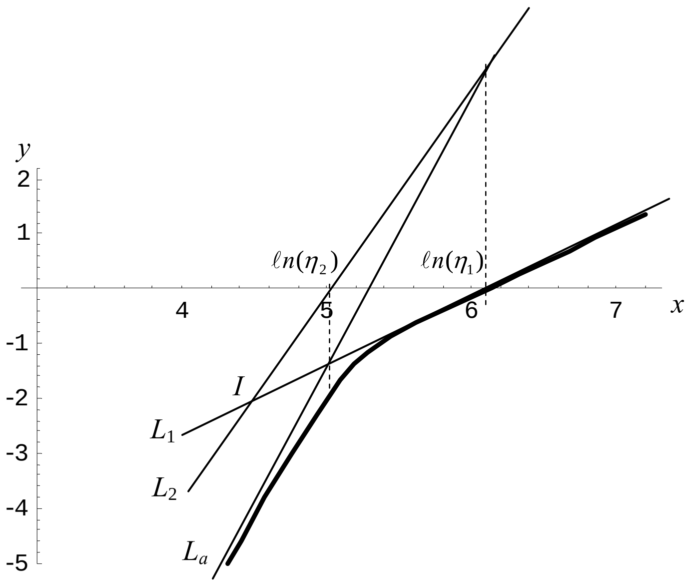


Fig. 5. $\beta_1 = 1.25,\ \eta_1 = 454.9;\ \beta_2 = 3.7,\ \eta_2 = 154.9 \qquad (\eta_1 >> \eta_2)$

Case (*iii*): $\eta_1 \cong \eta_2$

In this case, there are

$$x_A \cong x_B = \ell n(\eta_2) \quad \text{and} \qquad y_A \cong y_B \cong -0.6731. \tag{40}$$

A typical plot of y *vs.* x is shown in Fig. 6 and where $x_A \approx x_B \approx x_I$.

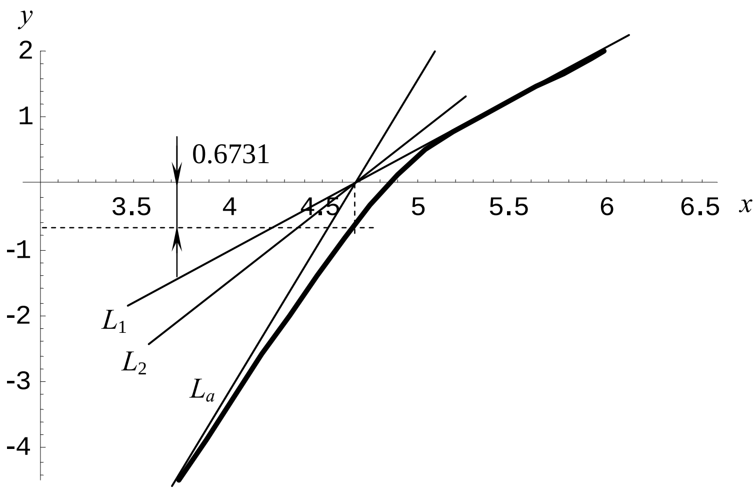


Fig. 6. $\beta_1 = 1.5,\ \beta_2 = 3.25,\ \eta_1 = \eta_2 = 105$

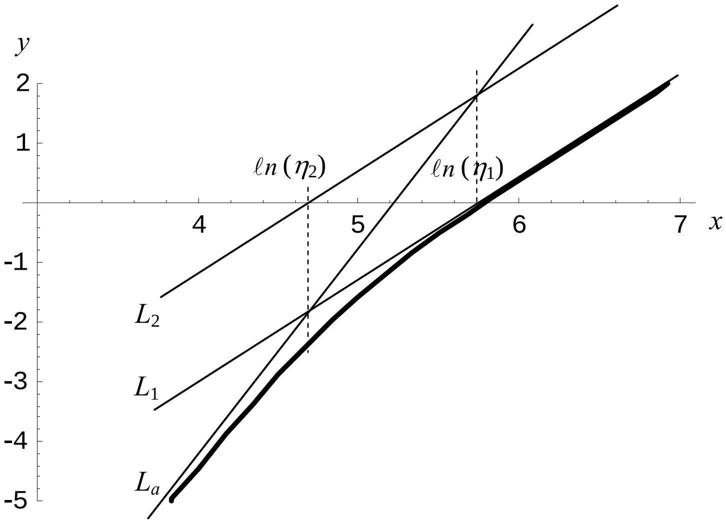


Fig. 7. $\beta_1 = \beta_2 = 1.75$, $\eta_1 = 326$, $\eta_2 = 105$

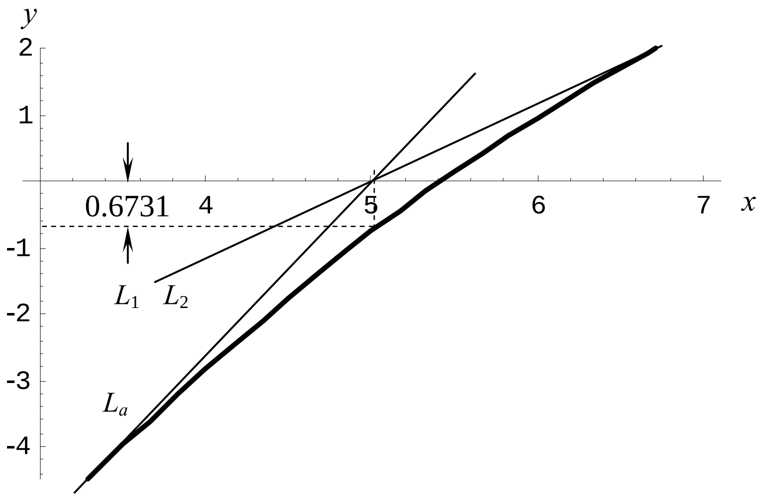


Fig. 8. $\beta_1 = \beta_2 = 1.25$, $\eta_1 = \eta_2 = 155$

3.1.2.2. Case of $\beta_1 = \beta_2$

When $\eta_1 \neq \eta_2$, the lines L_1 and L_2 are parallel to each other and a typical plot is as shown in Fig. 7. When $\eta_1 = \eta_2$, the lines L_1 and L_2 merge together and a typical plot is shown in Fig. 8.

Comment

Based on above analysis, it can be found that the 2-fold Weibull multiplicative model with $\eta_1 = \eta_2$ gives the special case of fitting plots that are modeled by exponentiated Weibull family model when $\theta > 1$. The exponentiated Weibull family is expressed by $F(t) = [F_w(t)]^\theta$ where $F_w(t) = 1 - \exp[-(t/\eta)^\beta]$ and θ is a parameter. The model property and parameter estimation are given in [6–8]. When $\theta = 2$, the exponentiated Weibull family model is a special case of the 2-fold Weibull multiplicative model.

3.2. Concave curve and associated models

3.2.1. Sectional model involving two Weibull distributions

3.2.1.1. Model-b

This model was discussed in Section 3.1.1 in detail. This model requires $\beta_1 \neq \beta_2$. If $\beta_1 < \beta_2$, the WPP plot is concave as shown in Fig. 9. The property of the fitting plot is the same as analyzed in Section 3.1.1 and the parameter estimates can be referred to [11].

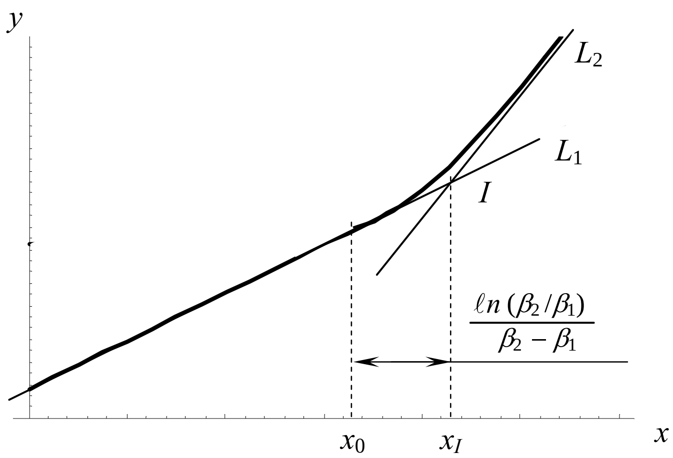


Fig. 9. Plot for Model-*b* ($\beta_1 < \beta_2$)

3.2.1.2. Model-c

Refer to Section 3.1.1 for the property of this model. This model requires $\beta_1 \neq \beta_2$. When $\beta_1 < \beta_2$, the model plot is a concave curve, see Fig. 10. Details for parameter estimates can be referred to [12].

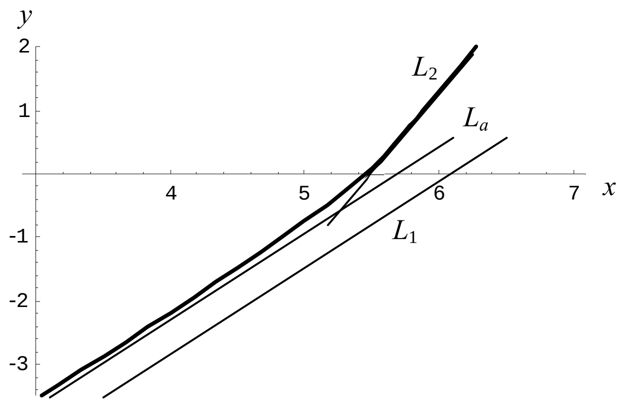


Fig. 10. WPP plot for $\beta_1 < \beta_2$ ($\beta > 1$)

$\beta_1 = 1.35$, $\eta_1 = 441$; $\beta_2 = 2.50$, $\eta_2 = 241$; $t_0 = 292.268$, $k = 1.837$.

3.2.2. Competing risk model

Competing risk model is applied to modeling an item's failure that is caused by more than one failure mode or cause, and these failure modes are statistically independent. The item fails whenever any failure mode occurs. The model is given by

$$R(t) = \prod_{i=1}^{n} R_i(t) \tag{41}$$

where, $R_i(t) = 1 - F_i(t)$, $F_i(t)$ is failure probability due to the *ith* failure mode or cause. Here, we are interested in the model as shown below:

$$R(t) = R_1(t) \cdot R_2(t) \tag{42}$$

with $R_i(\mathrm{t}) = exp[-(t/\eta_i)^{\beta_i}]$, $t \geq 0$. This model is analyzed in [11]. It needs $\beta_1 \neq \beta_2$, otherwise, the model reduces to a single Weibull distribution. Without loss of generality, it is assumed that β_1 is less than β_2.

Under the Weibull transformation given in equation (3), there is the following from equation (42),

$$y = y(x) = \beta_1 [x - \ell n(\eta_1)] + \ell n [1 + (\eta_1^{\beta_1} / \eta_2^{\beta_2}) e^{(\beta_2 - \beta_1)x}], \tag{43}$$

or

$$y = y(x) = \beta_2 [x - \ell n(\eta_2)] + \ell n [1 + (\eta_2^{\beta_2} / \eta_1^{\beta_1}) e^{-(\beta_2 - \beta_1)x}]. \tag{44}$$

This defines $y(x)$ is a non-linear function of x. Since $\beta_1 < \beta_2$, it is not difficult to obtain the followings:

$$\lim_{x\to-\infty} \ell n\,[1 + (\eta_1^{\beta_1} / \eta_2^{\beta_2})\, e^{(\beta_2-\beta_1)x}] = 0 \tag{45}$$

from equation (43) and

$$\lim_{x\to\infty} \ell n\,[1 + (\eta_2^{\beta_2} / \eta_1^{\beta_1})\, e^{-(\beta_2-\beta_1)x}] = 0 \tag{46}$$

from equation (44). That means C_f has two asymptotes L_1 as $x \to -\infty$ and L_2 as $x \to \infty$. C_f is concave as shown in Fig. 11.

Note that at $x = x_I$, $R_1(x_I) = R_2(x_I)$, where $I\,(x_I, y_I)$ is the intersection of L_1 & L_2, and as a result we have

$$y(x)\big|_{x=x_I} = y_I + \ell n(2) \tag{47}$$

and

$$dy(x)/dx\big|_{x=x_I} = (\beta_1 + \beta_2)/2. \tag{48}$$

Note that if $\beta_1 = 1$ in equation (42), this model reduces to a special case named as B distribution. B distribution has been applied to lifetime data analysis of systems where the system failure involves random failure and wear-out failure [37].

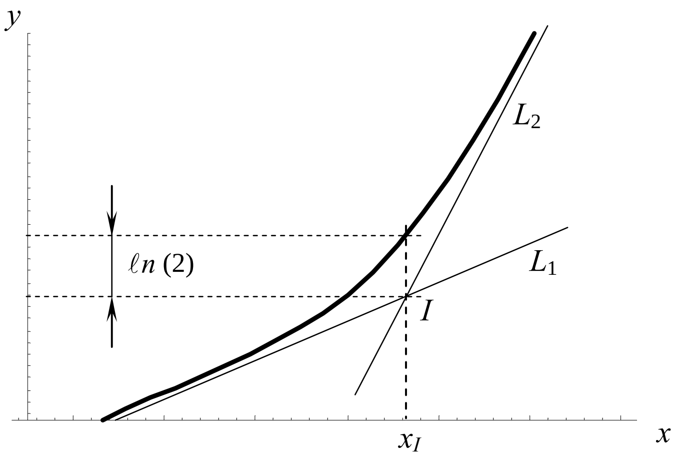


Fig. 11. The typical WPP plot for the competing risk model given by equation (42)

3.2.3. Sectional model involving three Weibull distributions

3.2.3.1. Model-d

Here, reliability function $R(t)$ is characterized by the following equations:

$$R(t) = \begin{cases} R_0(t), & 0 \le t \le t_0; \\ R_1(t)\cdot R_2(t), & t_0 < t < \infty. \end{cases} \tag{49}$$

Where, $R_i(t) = 1 - exp[-(t/\eta_i)^{\beta_i}]$, η_i and $\beta_i > 0$, $i = 0, 1, 2$.

According to the continuity of $R(t)$ and pdf at $t = t_0$, the parameters are constrained to satisfy the following equations:

$$(t_0/\eta_0)^{\beta_0} = (t_0/\eta_1)^{\beta_1} + (t_0/\eta_2)^{\beta_2}, \tag{50}$$

$$\beta_0\,(t_0/\eta_0)^{\beta_0} = \beta_1\,(t_0/\eta_1)^{\beta_1} + \beta_2\,(t_0/\eta_2)^{\beta_2}. \tag{51}$$

From equations (50) and (51), t_0 is obtained as

$$t_0 = \left[\frac{\beta_0 - \beta_2}{\beta_1 - \beta_0}\cdot\frac{\eta_1^{\beta_1}}{\eta_2^{\beta_2}}\right]^{1/(\beta_1-\beta_2)} = \left[\frac{\beta_2 - \beta_1}{\beta_2 - \beta_0}\cdot\frac{\eta_0^{\beta_0}}{\eta_1^{\beta_1}}\right]^{1/(\beta_0-\beta_1)}. \tag{52}$$

As $t_0 > 0$, it holds that either $(\beta_0 - \beta_2) > 0$ and $(\beta_1 - \beta_0) > 0$ or $(\beta_0 - \beta_2) < 0$ and $(\beta_1 - \beta_0) < 0$. Hence, $\beta_1 > \beta_0 > \beta_2$ or $\beta_1 < \beta_0 < \beta_2$. Without loss of generality, use the form $\beta_1 < \beta_0 < \beta_2$. If $\beta_1 = \beta_2$, this model reduces to a single distribution.

Under the Weibull transformation, equation (49) is transformed into

$$y = \beta_0\,[x - \ell n(\eta_0)], \quad -\infty < x \le \ell n(t_0); \tag{53}$$

and

$$y = \beta_1\,[x - \ell n(\eta_1)] + \ell n\left(1 + \frac{\eta_1^{\beta_1}}{\eta_2^{\beta_2}}e^{(\beta_2-\beta_1)x}\right) \tag{43}$$

or

$$y = \beta_2\,[x - \ell n(\eta_2)] + \ell n\left(1 + \frac{\eta_2^{\beta_2}}{\eta_1^{\beta_1}}e^{-(\beta_2-\beta_1)x}\right), \tag{44}$$

in the section $x_0 < x$. This implies C_f is a straight line for $-\infty < x \le \ell n(t_0)$, which is indicated by L_0 and then a smooth curve for $\ell n(t_0) < x < \infty$; see Fig. 12.

As discussed in Section 3.2.2, there are equations (45) and (46) so that C_f is a concave curve. As a result, the asymptotic behavior of equation (43) in its tendency with decreasing of x is given by

$$y = \beta_1\,[x - \ell n(\eta_1)]$$

as $x \to -\infty$. This is identical with L_1. Furthermore, from equation (44), the asymptotic behavior of the fitting plot is characterized by L_2:

$$y = \beta_2 [x - \ell n(\eta_2)]$$

as $x \to \infty$. Hence L_2 is the asymptote to the right end of C_f when $x \to \infty$. Similarly as before, let the intersection of L_1 and L_0 be II (x_{II} , y_{II}), there exist

$$x_{II} = [\,\beta_0\ \ell n(\eta_0) - \beta_1\ \ell n(\eta_1)\,] \,/\, (\,\beta_0 - \beta_1), \tag{54}$$

$$y_{II} = [\,\beta_0\ \beta_1\ \ell n(\eta_0 / \eta_1)\,] \,/\, (\,\beta_0 - \beta_1)\ . \tag{55}$$

Refer to Section 3.2.2, equations (47) and (48) apply for the model.

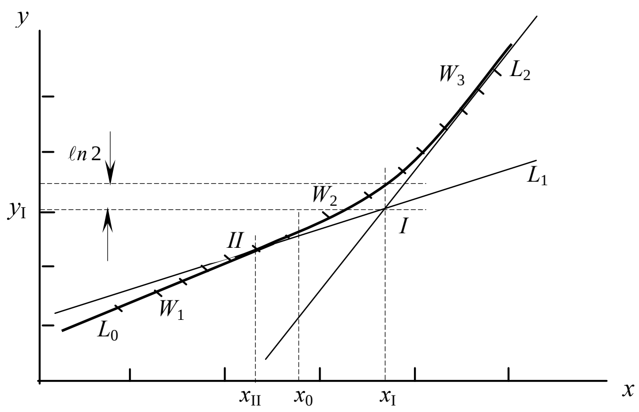


Fig. 12. Fitting plot for Model-*d*

Based on the model property, we can carry out parameter estimation as follows.

Parameter estimation

Plotting data on WPP is the same as what are given in Section 2.

#4. Use a straight line to fit the left side of the plot, the slope of this line yields β_0 and the x-axis intercept gives $\ell n\,(\eta_0)$.

#5. Try to draw the asymptote L_2 to the right side of the plot with increase of x. Its slope gives β_2 and the x-axis intercept yields $\ell n\,(\eta_2)$.

#6: To determine the point I on L_2, from which the vertical distance to C_f is $\ell n(2)$. The horizontal coordinate of this point yields x_I. Draw a tangent to C_f at $x = x_I$, its slope gives $(\beta_1 + \beta_2)/2$. Thus β_1 is obtained. Draw a straight line through this point with slope β_1, as a result, L_1 is determined. The x-axis intercept of L_1 yields $\ell n\,(\eta_1)$.

#7. Calculate t_0 by using equation (52) and hence $x_0 = \ell n(t_0)$ is known.

If the calculated x_0 is not identical with what should be through careful observation, readjust a little bit L_2 and further L_1, and repeat Steps 5 ~ 7.

3.2.3.2. Model-e

Similar to Model-*d*, $R(t)$ is given by

$$R(t) = \begin{cases} R_1(t) \cdot R_2(t), & 0 \le t \le t_0; \\ R_0(t), & t_0 < t < \infty. \end{cases} \tag{56}$$

Where, $R_i(t)$ is reliability function of Weibull distribution with parameters β_i and η_i (i = 1, 2).

The character of fitting plot for this model is shown in Fig. 13. The plot is concave and its right side with larger x is a straight line which is identical with L_0 given by $y = \beta_0 [x - \ell n(\eta_0)]$.

Refer to the procedure given for Model-*d*, the parameters of this model can be estimated in a similar way.

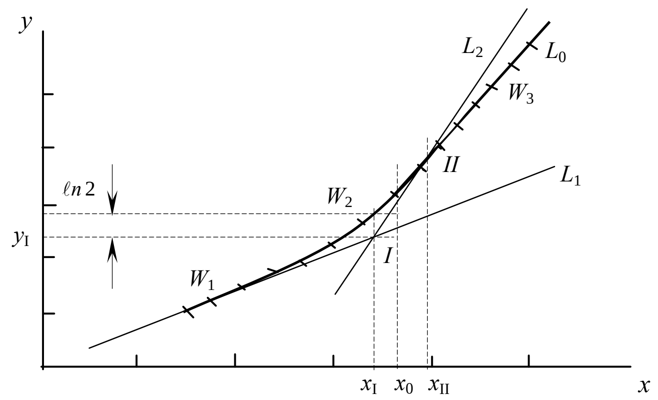


Fig. 13. Character of fitting plot of Model-*e* on WPP

Comment

The shape of fitting plot determined by Model-*d* or Model-*e* is similar to the competing risk model shown in Section 3.2.2. If one fitting plot looks like to take the shape character of these models. The competing risk model is first applied and then justify if the condition $y(x) = y_I + \ell n\, 2$ at $x = x_I$ is satisfied and the calculated value of $dy(x)/dx$ at $x = x_I$ is coincident with what looks like in the figure. If either of them is satisfied, the competing risk model is a good choice. Otherwise, try to use Model-*d* and Model-*e*. It is not difficult to choose one from them. Or, if the right side of the plot looks quite like on a straight line, Model-*c* as given in Section 3.2.1.2 is also a good choice. Else, if the left side (where the data are smaller) of the plot looks quite like on a straight line, Model-*b* given in Section 3.2.1.1 is appropriate for modeling the given set of data.

3.3. Another sectional model involving two Weibull distributions —Model-f

Model-*f* is a sectional model, it is

$$F(t) = \begin{cases} kF_1(t), & 0 \le t \le t_0; \\ 1 - kR_2(t), & t_0 < t < \infty. \end{cases} \tag{57}$$

Where, k is a parameter, $k > 0$; $R_i(t)$ is reliability function of the Weibull distribution with parameters β_i and η_i (i = 1, 2) and $F_1(t) = 1 - R_1(t)$. Note that this model is the same as Model-*c* when $t \in [0, t_0]$ and Model-*b* when $t \in (t_0, \infty)$.

It is similar to Section 3.1.1, by using the Weibull transform, the model plot on WPP is obtained as

$$y = y(t) = \begin{cases} \ell n\,\{-\ell n\,[1 - k + k\,exp(-(e^x/\eta_1)^{\beta_1})]\}, & -\infty < x \le \ell n(t_0) \quad (58) \\ \beta_2[\,x - \ell n(\eta_2)\,] + \ell n\,[1 - \eta_2^{\beta_2}\ell n(k)e^{-\beta_2 x}], & \ell n(t_0) < x < \infty. \quad (59) \end{cases}$$

Note that equation (58) is the same as equation (29) for $x \in (-\infty, \ell n(t_0)]$ and equation (59) is the same as equation (20) for $x \in (\ell n(t_0), \infty)$. From Section 3.1.1, it is known that the plot of equation (58) is convex when $k < 1$ and concave when $k > 1$, whereas the plot of equation (59) is concave when $k < 1$ and convex when $k > 1$. Therefore, the model plot is S-shaped or inversely S-shaped, and there is one inflection.

According to the discussions given in Section 3.1.1 and as analyzed in [12–13], it is concluded as below in three cases.

Case 1: $\beta_1 = \beta_2$

If $c < 1$ ($c = (\eta_1/\eta_2)^{\beta_2}$), then $k < 1$; if $c > 1$, then $k > 1$.

If $c < 1$, then $\eta_1 < \eta_2$; if $c > 1$, then $\eta_1 > \eta_2$.

Case 2: $\beta_1 < \beta_2$

$k > 1$ or $k < 1$.

Case 3: $\beta_1 > \beta_2$

$k > 1$ or $k < 1$.

Hence, the property of the model plot on WPP is determined by k as follows:

When $k > 1$, it is concave for $x \in (-\infty, \ \ell n(t_0))$ and then convex for $x > \ell n(t_0)$;

when $k < 1$, it is convex for $x \in (-\infty, \ \ell n(t_0))$ and then concave for $x > \ell n(t_0)$.

The model plot shows S-shaped or inversely S-shaped curve, and there is one inflection. If assuming this inflection as I_f, the x coordinate of I_f, x_{If}, is given by $\ell n(t_0)$, *i.e.*, $x_{If} = \ell n(t_0)$. As an illustration, the model plots are shown in Fig. 14 for the two cases of $k > 1$ and $k < 1$.

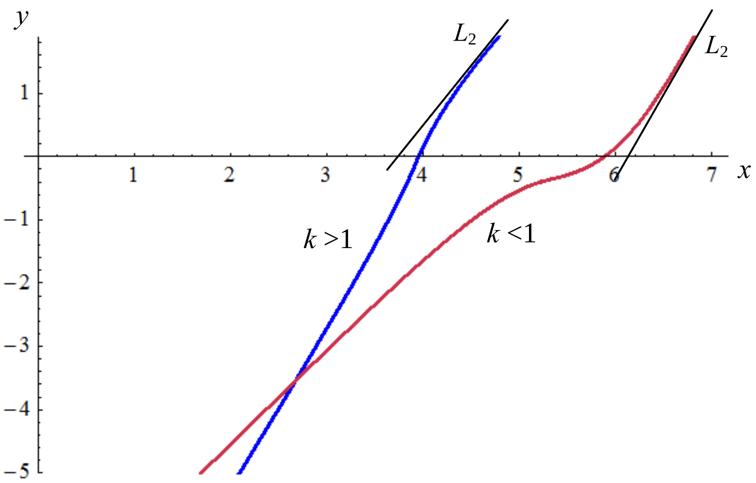


Fig. 14. Typical plots of Model-*f* on WPP

Case $k < 1$: $\beta_1 = 1.5$, $\eta_1 = 100$, $\beta_2 = 3.0$, $\eta_2 = 5\eta_1$, $k = 0.5136$, $t_0 = 211.910$;

Case $k > 1$: $\beta_1 = 2.5$, $\eta_1 = 100$, $\beta_2 = 1.5$, $\eta_2 = 0.3\eta_1$, $k = 3.593$, $t_0 = 50.177$.

3.4. Mixed Weibull distribution

A mixed distribution is given by

$$f(t) = \sum_{i=1}^{n} p_i \cdot f(t_i) \tag{60}$$

with $\sum_{i=1}^{n} p_i = 1$ and where $p_i > 0$, $f_i(t)$ ($i = 1$ to n) stands for probability density function of subpopulation i.

Here, we are interested in the mixed Weibull distribution given by

$$f(t) = p \cdot f_1(t) + q \cdot f_2(t) \tag{61}$$

where, p is mixing weight, $p \in (0, 1)$, $p + q = 1$ and $f_i(t) = \beta_i \eta_i^{-\beta_i} t^{\beta_i - 1} exp[-(t/\eta_i)^{\beta_i}]$, $t \geq 0$, $i = 1, 2$. Without loss of generality, we assume that $\beta_1 < \beta_2$. If $\beta_1 = \beta_2$, it needs $\eta_1 \neq \eta_2$ and we assume $\eta_1 >$

η_2 without loss of generality; otherwise, it reduces to one Weibull distribution. The detailed analysis of the model and parameter estimation can be referred to [9,10,31–33,38]. From equation (61), it is derived that

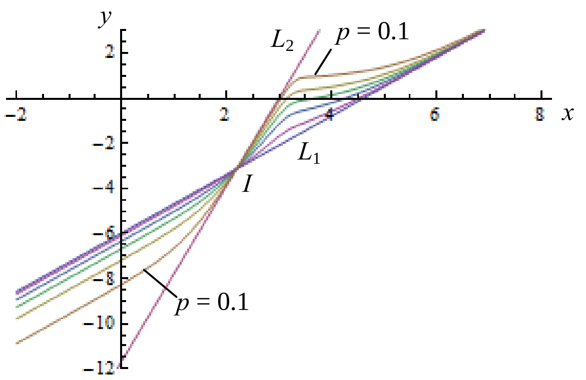
$$R(t) = p \cdot R_1(t) + q \cdot R_2(t). \tag{62}$$

The plot of the model on WPP is as follows:

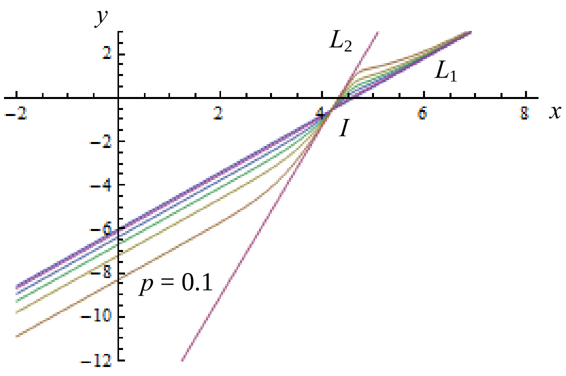
$$y = y(x) = \ell n\{-\ell n[pR_1(e^x) + qR_2(e^x)]\}. \tag{63}$$

The plot of equation (63) is constrained by the two straight lines L_1 and L_2, see Figs. 15 and 16. If $\beta_1 \neq \beta_2$, L_1 and L_2 has an intersection, I, which is the first inflection of the plot of equation (63). In fact, the plot has two inflections as shown in Fig. 15. If $\eta_2 < \eta_1$, I locates below x-axis. If $\eta_2 > \eta_1$, I locates above x-axis. I moves up and right with increasing of η_2 giving L_1. If x tends to be very large, the plot tends to an asymptote defined by L_1. On the other hand, if x becomes very smaller, the plot tends to another asymptote which is parallel to L_1. The distance from this asymptote to L_1 is determined by the value of p. That is, the smaller the value of p, the larger the distance.

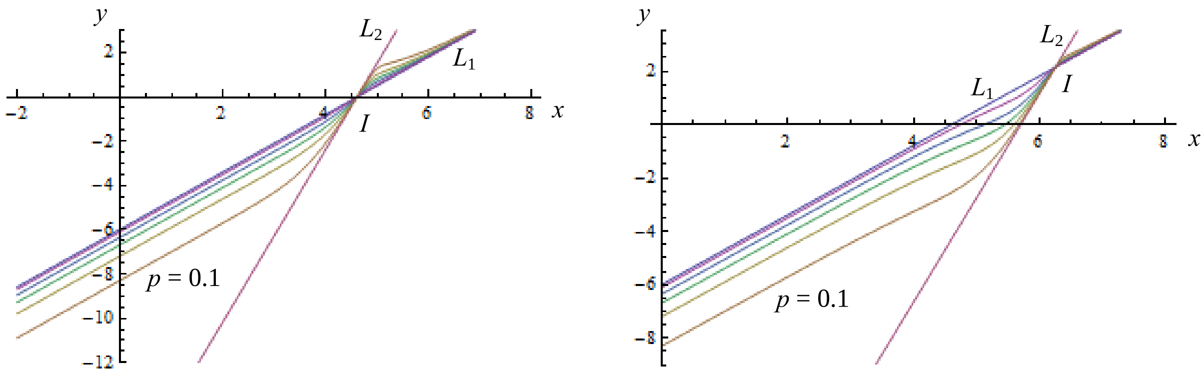
If $\beta_1 = \beta_2$, L_1 is parallel to L_2 and the plot of equation (63) has only one inflection, see Fig. 16. When x becomes very large, the plot tends to L_1 which is an asymptote. If x becomes quite smaller, there is another asymptote to the left side of the plot of equation (63), which is parallel to L_1 and L_2. The distance between this asymptote and L_2 becomes smaller with decreasing of p. These properties are important in model selection and parameter estimation.



(a) $\eta_2 = 0.2\eta_1$



(b) $\eta_2 = 0.75\eta_1$



(*c*) $\eta_2 = \eta_1$ (*d*) $\eta_2 = 3\eta_1$

Fig. 15. Character of the mixed Weibull model plot on WPP
$\beta_1 = 1.3$, $\beta_2 = 3.9$, $\eta_1 = 100$, $p = 0.1, 0.3, 0.5, 0.7, 0.9$

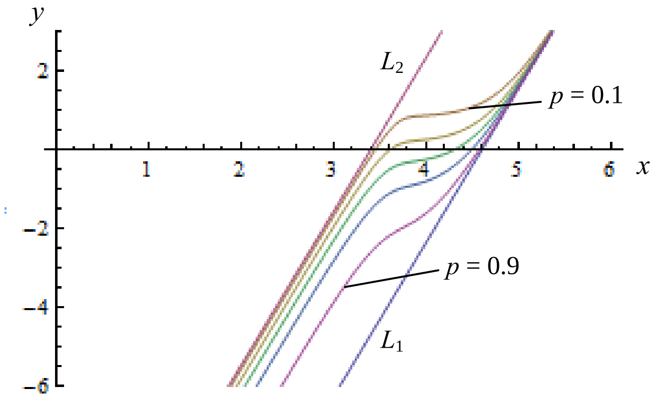


Fig. 16. Character of the mixed Weibull plot on WPP
$\beta_1 = \beta_2 = 3.9$, $\eta_1 = 100$, $\eta_2 = 0.3\eta_1$, $p = 0.1, 0.3, 0.5, 0.7, 0.9$

4. Basic procedure for choosing an optimal model

Selecting an appropriate model is important for analysis of a given set of test data. One can use two different ways to choose a model for the given data set. One is through analysis of physics of failure and the other is by examining the shape properties of the data plot on WPP. We discuss these two approaches in this section with focus on model selection based on the data plot on WPP.

4.1. Choose a model by analysis of physics of failure

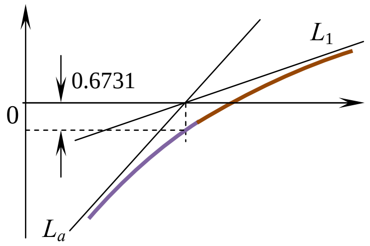
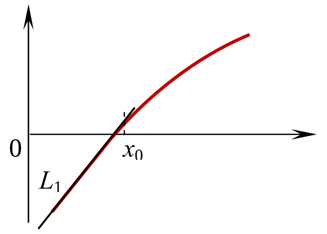
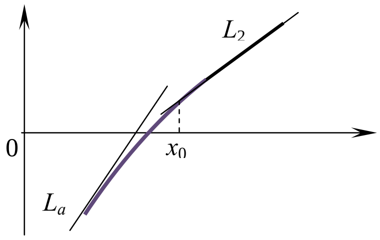
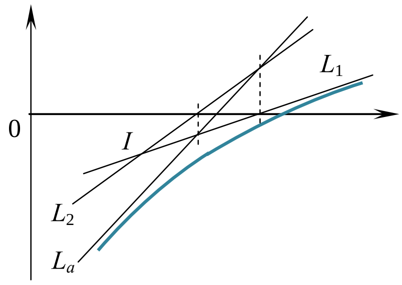
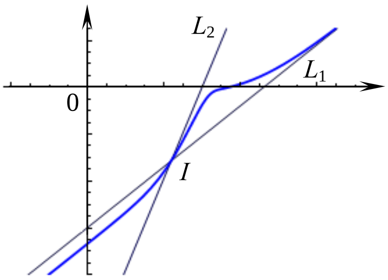
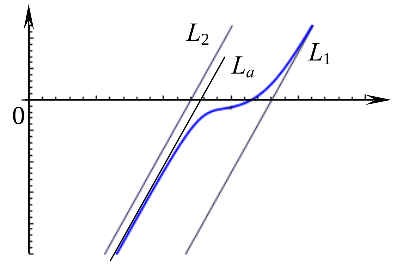
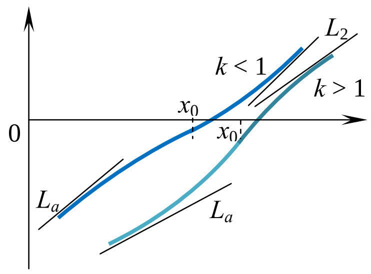
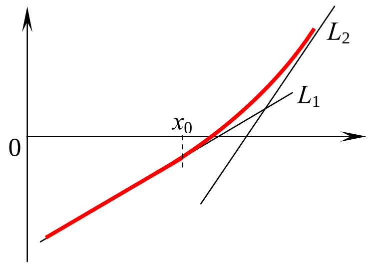
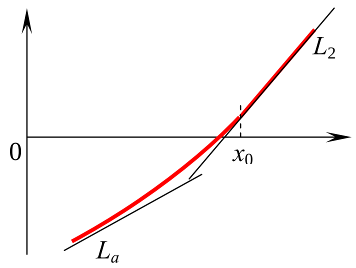
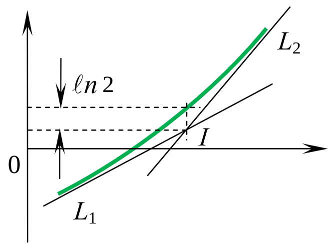
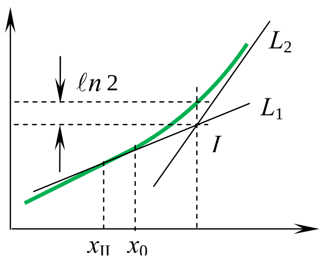
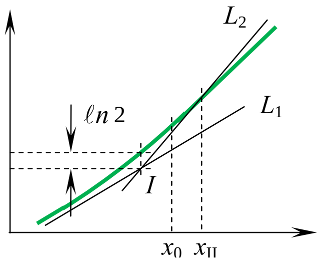
To select an appropriate model, the following aspects need to be considered:

1. What is the failure mode or dominant failure modes in various life periods?
2. What is the failure distribution of the similar products in history?
3. Is an item failure caused by two or more causes or modes which are independent of each other? If yes, one can try the competing risk model.
4. Is the time-to-failure of an item or system represented by the largest time-to-failure of components that consist of the system? If yes, one may try multiplicative model.
5. Can the given data set be divided into subgroups and the data in one subgroup be treated as coming from one subpopulation? If yes, one can try mixed distribution model. If a system's failure is caused by more than one cause, a mixed Weibull model would be appropriate to use in modeling the lifetime of the system.
6. Does an item's failure behavior show difference in various life periods? If yes, the sectional distribution models may be appropriate for modeling.

4.2. Choose a model according to the shape property of the plot

An easy and intuitive way to model a given set of data is by the property of the fitting plot of the data on WPP. As discussed in Section 3, each model gives its special property for the model plots on WPP. Therefore, based on these properties, we can easily choose an appropriate one for modeling the given data set, see Table 2 for model selection.

Table 2 Property of model plots

		
Multiplicattive Weibull model with ($\eta_1 = \eta_2$)	Model-*a* Model-*b* ($\beta_1 > \beta_2$)	Model-*c* ($\beta_1 > \beta_2$)
		
Multiplicative Weibull model	Mixed Weibull model ($\beta_1 \neq \beta_2$)	Mixed Weibull model ($\beta_1 = \beta_2$)
		
Model-*f* with $k > 1$ and $k < 1$	Model-*b* ($\beta_1 < \beta_2$)	Model-*c* ($\beta_1 < \beta_2$)
		
Competing risk model	Model-*d*	Model-*e*

In general, the procedure for choosing an optimal model by the graphical approach is as follows.

a. Plot data on WPP.

b. Analyze the shape property of the plot.

c. Use Table 2 and choose a model for the plot according to its shape property. Some times, several models could be appropriate.

d. Perform parameter estimates of the selected models through combination with the Least Squares method.

e. Make comparison of R^2 values between different models selected or carry out goodness-of-fit test, or calculate AIC (Akaike's Information Criterion) values [39] through MLE to evaluate the selected models and then to choose a best one for the given set of data.

4.3. Models involving three Weibull distributions

A few other forms of Weibull models involving three Weibull distributions were discussed in [13] and [35]. For a complex data set, if the Weibull models with three parameters and all the models as discussed in this paper are not satisfied for modeling, one may have a trial to construct other forms of sectional models involving multiple Weibull distributions.

In this section, a general procedure for choosing an optimal model for a given set of data is proposed. In order to give an illustration, two examples are presented in the following section.

5. Examples

Example 1

A set of data is shown in Table 3. Each of the data represents the time until death of mile mice exposed to 300 rads of radiation. This group of male mice was maintained in a germ-free environment. Though this group of data represents the survival days of male mice due to thymic lymphoma, the failure could also be involved in the effect of other failure modes and other causes. It is of interest, for example, to evaluate the effect of a germ-free environment on the incidence rate of reticulum cell sarcoma while accommodating the competing risks of developing thymic lymphoma or other causes of failure [40]. This set of data is selected for demonstrating the data modeling method as discussed in this present paper.

Table 3 Days of until death of male mice exposed to 300 rads of radiation (Germ-Free Group) [40]

158	192	193	194	195	202	212	215	229	230
237	240	244	247	259	300	301	321	337	415
434	444	485	496	529	537	624	707	800	

Based on the procedure as described in Section 3, plot the data on WPP as shown in Fig. 17. Clearly the data points do not scatter on a straight line and the data plot shows first convex and then concave. It is clearly to show that the fitted plot has one inflection. Therefore, the mixed Weibull model and Model-*f* may be appropriate to model this set of test data according to Table 2. After trial, however, it is found that the mixed Weibull distribution does not fit to modeling this set of data. Model-*f* is an appropriate one for this set of data. Finally, the model parameters are obtained with corresponding to calculation of R^2 value for the model and associated parameters. The estimated parameters are β_1 = 8.25, η_1 = 234.25, β_2 = 2.59, η_2 = 506.56 and k = 0.556 with R^2 = 0.992. That means this model fits the data set very well. See the model fitting and the data plot in Fig. 17. The survival probability and the empirical plot are presented in Fig. 18. Again, it shows that the model fits the data quite well.

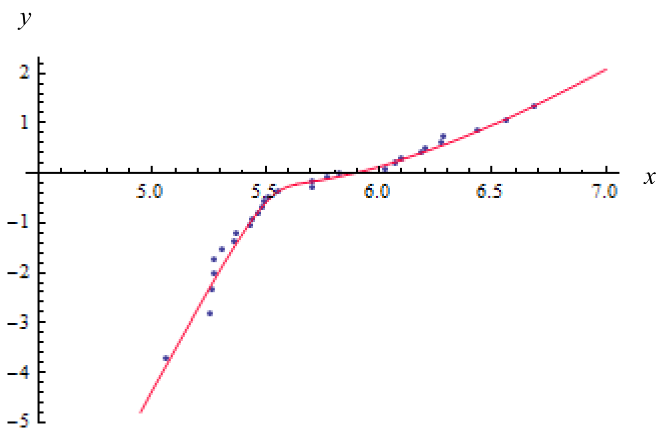


Fig. 17.   Data plot (dot line) and model fitting (solid curve)
β_1 = 8.25, η_1=234.25; β_2=2.59, η_2 =506.60; k = 0.556.

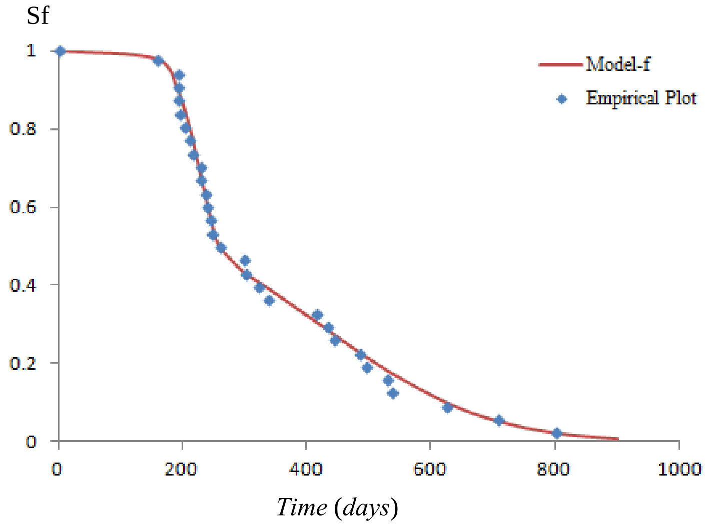


Fig. 18. Survival probability curves of the empirical distribution and the fitted distribution of Model-*f* ($R^2 > 0.99$)

Example 2

A sample of data representing months of survival of 26 patients being treated for one disease is shown in Table 4, of whom 13 patients taking irinotecan plus cisplatin and another 13 patients taking etoposide plus cisplatin. This set of data was presented for illustration of data processing and analysis in [41]. As described in Section 3, the WPP plot of this set of data is generated first as shown in Fig. 19. By observing carefully the characterization of this plot, one can judge that Model-*a* and the multiplicative Weibull model would be appropriate ones for fitting this set of data. After trial, it is found that Model-*a* is better than the multiplicative Weibull model. At the same time, another potential 3-parameter Weibul model proposed by Dimitrakopoulou *et al.* [42] was also selected for verification. Here, it is named as D-K-S model. This model is given by $R(t)$ = $\exp\{1-(1+\lambda t^{\beta})^{\alpha}\}$ for $t > 0$, where λ is scale parameter, α and β are of shape parameter (λ, α, $\beta > 0$).

By fitting the data plot using Model-*a* and D-K-S model, it is found that Model-*a* is better than D-K-S model for modeling this set of data. The parameters of D-K-S model are obtained by the method of maximum likelihood estimate, which are $\alpha = 0.3174$, $\beta = 2.650$ and $\lambda = 0.007758$. And the R^2 value with the estimated parameters is 0.97364. The parameters of Model-*a* are estimated by the graphical approach, which are $\beta_1 = 2.078$, η_1=12.897; β_2=0.380, η_2 =1.380 and $\gamma = 11.930$. And the R^2 value for Model-*a* with the estimated parameters is 0.9902. Therefore, Model-*a* would be an optimal model for modeling this set of data. The fitting plot of Model-*a* is shown in Fig. 19 and the survival probability curves of the empirical distribution given in dotted line, and Model-*a* and D-K-S model fitted with the estimated parameters are shown in Fig. 20. It is clearly verified that Model-*a* fits the data very well.

Table 4 Months of survival of patients after treatment [41]

13.57	11.7	12.52	30.65	2.73	25.49	13.31	16.89	10.94
8.18	9.72	15.61	56.38	8.11	5.82	1.94	13.34	7.56
6.47	14.69	16.26	9.43	9.49	4.86	14.65	5.95	

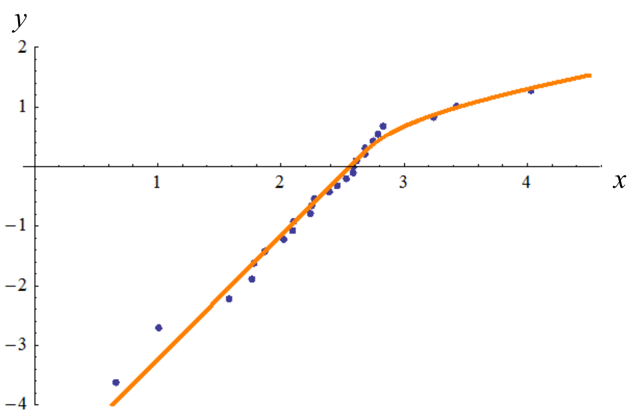


Fig. 19. Data plot and fitting plot of Model-*a* with

β_1 = 2.078, η_1=12.897; β_2=0.380, η_2=1.380; γ = 11.930; R^2 = 0.9902.

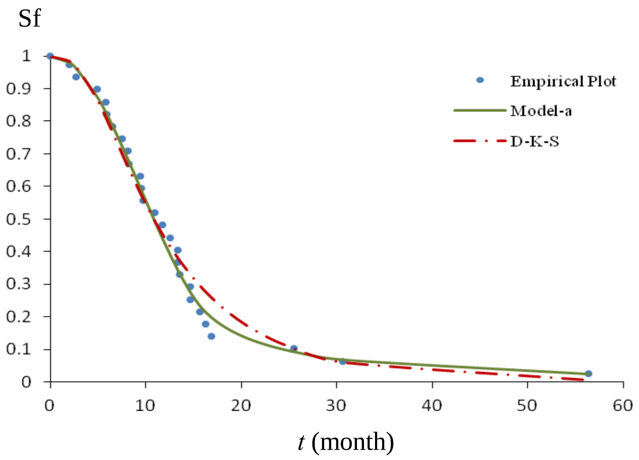


Fig. 20. Survival probability curves of the empirical distribution and the fitted distribution of Model-*a* and D-K-S model

6. Remarks

In this paper, the property of Weibull-related models involving more than three parameters is summarized and the method and procedure for choosing an optimal model to formulate a given set of data are proposed. The graphical approach can provide us with shape characterization which a failure data set has, so that it yields some insight for us to select a better model to fit the data. This is straightforward and easy to be applied. It can provide the initial estimates of model parameters. Although the parameter estimates by graphical representation are approximate to some extent, the initial estimates can be refined through combination with the method of Least Squares.

In the graphical method for model parameter estimates, the asymptotic property of the fitting plot of the given data set is frequently utilized. If the model function converges very fast to its limit when $x \to \infty$, or $x \to -\infty$, it is easy to draw an asymptote with higher accuracy. Otherwise, it will involve error when to fit a curve end with a straight line. To fit an accurate asymptote to the end part of a section of the plot is very important in parameter estimates. Often one's expertise in parameter estimates by the graphical approach plays an important role in improving accuracy of the estimation.

After a few appropriate models or an optimal model is determined, the next step is to perform parameter estimation of the selected models. As parameter estimation by the graphical approach yields an initial estimate and it involves a certain degree of subjectivity, other more accurate statistical methods are necessary to be applied, such as MLE, Least Squares estimation, *etc*. For MLE method for the parameter estimates of the sectional models, the recursive method should be applied as the section partition points (such as t_0 and t_1) can not be chosen accurately from the plot. In general, it might require several iterations before a good fit is obtained. Often is it to apply the recursive method to solving a set of equations that have several parameters involved.

Acknowledgement

The authors are grateful to the anonymous reviewers for their insightful comments and constructive suggestions that have greatly improved the paper.

REFERENCES

[1] Lai CD, Xie M and Murthy DNP. A Modified Weibull Distribution. IEEE Transactions on Reliability 2003;52(1):33–37.

[2] Xie M, Tang Y, Goh TN. A modified Weibull extension with bathtub-shaped failure rate function. Reliability Engineering and System Safety 2002;76(3):279–85.

[3] Marshall AW, Olkin I. A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families. Biometrika 1997;84:641–52.

[4] Bebbington M, Lai CD, Zitikis R. A flexible Weibull extension. Reliability Engineering and System Safety 2007;92:719–26.

[5] Zhang T, Xie M. Failure data analysis with extended Weibull distribution. Communications in Statistics—Simulation and Computation 2007;36(3):579–92.

[6] Mudholkar GS, Srivastava DK. Exponentiated Weibull family for analyzing bathtub failure-rate data. IEEE Trans. Reliability 1993;42:299–302.

[7] Mudholkar GS, Srivastava DK, Freimer M. Exponentiated Weibull family: A re-analysis of the Bus-Motor-Failure data. Technometrics 1995;37(4):436–45.

[8] Jiang R, Murthy DNP. The exponentiated Weibull family: A graphical approach. IEEE Trans. Reliability 1999;48(1):68 –72.

[9] Jiang S, Kececioglu D. Graphical representation of two mixed-Weibull distributions. IEEE Transactions on Reliability 1992;41(2):241–47.

[10] Jiang R, Murthy DNP. Modeling failure-data by mixture of 2 Weibull distributions: A graphical approach. IEEE Trans. Reliability 1995;44(3):477–88.

[11] Jiang R, Murthy DNP. Reliability modeling involving two Weibull distributions. Reliability Engineering and Systems Safety 1995;47(3):187–98.

[12] Jiang R, Zuo MJ, Murthy DNP. Two sectional models involving two Weibull distributions. International Journal of Reliability, Quality and Reliability Engineering 1999;6(2):103–22.

[13] Jiang R, Murthy DNP. Two sectional models involving three Weibull distributions. Quality and Reliability Engineering International 1997;13(2):83–96.

[14] Murthy DNP, Jiang R. Parametric study of sectional models involving two Weibull distributions. Reliability Engineering and System Safety 1997;56(2):151–59.

[15] Murthy DNP, Xie M, Jiang R. Weibull Models. New York: John Wiley & Sons, 2003.

[16] Gupta RC, Lvin S and Peng C. Estimating turning points of the failure rate of the extended Weibull distribution. Journal of Computational Statistics and Data Analysis 2010;54(4):924–34.

[17] Cordeiro GM and Lemonte AJ. On the Marshall–Olkin extended Weibull distribution. Statistical Papers 2012;(53). DOI 10.1007/s00362-012-0431–8.

[18] Pappas V, Adamidis K and Loukas S. A family of lifetime distributions. International Journal of Quality, Statistics, and Reliability 2012. doi:10.1155/2012/760687.

[19] Carrasco JMF, Ortega EMM and Cordeiro GM. A generalized modified Weibull distribution for lifetime modeling. Computational Statistics & Data Analysis 2008;53(2):450–62.

[20] Silva GO, Ortega EMM and Cordeiro GM. The beta modified Weibull distribution. Lifetime Data Analysis 2010:16(3):409–30.

[21] Gasmi S. and Berzig M. Parameters estimation of the modified Weibull distribution based on Type I censored samples. Applied Mathematical Sciences 2011;5(59):2899–2917.

[22] Almalki SJ and Yuan JS. A new modified Weibull distribution. Reliability Engineering and System Safety 2013;111. 10.1016/j.bbr.2011.03.031.

[23] Murthy DNP, Bulmer M and Eccleston JA. Weibull model selection for reliability modeling. Reliability Engineering and Systems Safety 2004;86:257–67.

[24] Lai CD, Murthy DNP and Xie M. Weibull distributions, *Wiley Interdisciplinary Reviews*: *Computational Statistics*, Vol. 3 (2011), No. 3, pp. 282–7.

[25] Mann NR, Schafer RE, Singpurwalla ND. Methods for Statistical Analysis of Reliability and Life Data. New York: John Wiley & Sons, 1974.

[26] Goetghebeur E, Ryan L. Analysis of competing risks survival data when some failure types are missing. Biometrika 1995;82(4):821–33.

[27] Pintilie M. Competing risks: A practical perspective. John Wiley & Sons, 2006.

[28] Stephenson J, Chadwick BL, Playle RA, Treasure ET. A competing risk survival analysis model to assess the efficacy of filling carious primary teeth. Caries Research 2010;44(3):285–93.

[29] Ma Z, Krings AW. Competing risks analysis of reliability, survivability, and prognostic and health management. 2008 IEEE Aerospace Conference, p. 1–21, March 1 ~ 8, 2008.

[30] ReliaSoft, Multimodal analysis. Reliability HotWire 2002, No. 2. *http://weibull.com/hotwire/issue12/hottopics12.htm* (last assess in December 2012).

[31] Kececioglu D, Wang W. Parameter estimation for mixed-Weibull distribution. Proceedings of Annual Reliability and Maintainability Symposium 1998:247–52.

[32] Jiang S, Kececioglu D. Maximum likelihood estimates, from censored data, for mixed-Weibull distributions. IEEE Transactions on Reliability 1992;41(2):248–55.

[33] Ling D, Huang HZ, Liu Y. A method for parameter estimation of mixed Weibull distribution. Proceedings of Annual Reliability and Maintainability Symposium 2009:129–33.

[34] Elandt-Johnson RC, Johnson NL. Survival Model and Data Analysis. New York: John Wiley & Sons, p. 192–196, 1980.

[35] Zhang T, Ren Y. Failure data analysis by models involving 3 Weibull distributions. Annual Symposium on Reliability and Maintainability 2002; p. 44–50.

[36] Nelson W. Applied Life Data Analysis. New York: John Wiley & Sons, 1982.

[37] Bousquet N, Bertholon H, Celeux G. An alternative competing risk model to the Weibull distribution for modeling aging in lifetime data analysis. Lifetime Data Analysis 2006;12:481–504.

[38] Gong ZJ. Estimation of mixed Weibull distribution parameters using the SCEMUA algorithm: application and comparison with MLE in automotive reliability analysis. Reliability Engineering and System Safety 2006;91:915–22.

[39] Burnham KP, Anderson DR. Model Selection and Multimodel Inference—A Practical Information-Theoretic Approach. New York: Springer-Verlag, 2nd edition, 2002.

[40] Kalbfleisch JD, Prentice RL. The Statistical Analysis of Failure Time Data. New York: John Wiley & Sons, 2nd eition, 2002, p. 389–90.

[41] Dawson B, Trapp RG. Basic & Clinical Biostatistics. New York: Lange Medical Books/McGraw-Hill, Medical Publishing Division, 4th edition, 2004, p. 225.

[42] Dimitrakopoulou T, Adamidis K, Loukas S. A lifetime distribution with an upside-down bathtub-shaped hazard function. IEEE Transactions on Reliability 2007;56(2):308–11.