Appl. Math. J. Chinese Univ.
2016, 31(4): 425-438

Asymptotic properties and expectation-maximization algorithm for maximum likelihood estimates of the parameters from Weibull-Logarithmic model

GUI Wen-hao[1] ZHANG Huai-nian[2]

Abstract. In this article, we consider a lifetime distribution, the Weibull-Logarithmic distribution introduced by [6]. We investigate some new statistical characterizations and properties. We develop the maximum likelihood inference using EM algorithm. Asymptotic properties of the MLEs are obtained and extensive simulations are conducted to assess the performance of parameter estimation. A numerical example is used to illustrate the application.

§1 Introduction

The statistical modeling of life time distribution has important applications in various practical problems such as biological and engineering sciences. Various distributions have been introduced by mixing some useful life distributions and analyzed with respect to different characteristics. [1] compounded an exponential distribution with a geometric distribution and proposed an exponential geometric (EG) distribution which has decreasing failure rate function. [4] extended the model to a Weilbull-geometric (WG) distribution by replacing exponential distribution by a Weibull distribution. [14] studied an exponential logarithmic (EL) distribution and discussed its properties. In this article, we consider a Weibull logarithmic (WL) distribution which is obtained by mixing a Weibull distribution with a logarithmic distribution. This distribution was first introduced by [6] and is becoming increasingly popular for modeling lifetime data.

[6] provided some basic properties of density function, distribution function, moment generating function, uncertainty measures and presented an example to calculate uncertainty measures. In this paper, we focus on the asymptotic properties and expectation-maximization algorithm for maximum likelihood estimates of the parameters from this distribution.

Received: 2015-08-18.
MR Subject Classification: 62E15, 62F10.
Keywords: Maximum likelihood estimate, EM algorithm, Fisher information, Order statistics, Asymptotic properties.
Digital Object Identifier(DOI): 10.1007/s11766-016-3391-2.
Supported by the program for the Fundamental Research Funds for the Central Universities (2014RC042, 2015JBM109).

The rest of the paper is organized as follows: Section 2 studies the Weibull-Logarithmic distribution and provides its basic statistical properties. In Section 3, we study the sampling distributions of several extreme order statistics. Section 4 investigates the maximum likelihood inference based on EM algorithm and the asymptotic properties of the MLEs. We also conduct simulations in this Section. A real illustrative application is proposed in Section 5.

§2 Basic properties of Weibull-Logarithmic distribution

2.1 Model background

[6] obtained the Weibull-Logarithmic distribution by mixing Weibull and Logarithmic distribution. Here we further describe and explain the motivation as follows:

We consider a situation where failure of a unit occurs due to the presence of unknown Z initial defects of same kind. Let $Y_1, Y_2, \ldots, Y_Z$ be the corresponding lifetimes and each defect can be determined only after the failure, in which case it is repaired. Therefore, $X = \min(Y_1, \ldots, Y_Z)$ models the time of the first failure.

In practice, the known Weibull distribution has various shapes of failure rate function and is widely used for the inference of life information. Using Weibull distribution as a member of the compounding distribution is a good trial. The Logarithmic distribution is commonly used to fit species abundance data with long tails and it has many applications in various fields of research such as ecology, economics, biology etc. See [11] for more details. It is attractive for practitioners to take a compounding distribution for lifetime inference if its failure rate function has many shapes.

Assume the failure times $Y_1, Y_2, \ldots, Y_Z$ follow a Weibull distribution with the following probability density function:

$$f(y) = \alpha\beta y^{\alpha-1} e^{-\beta y^{\alpha}}, \quad y > 0, \tag{1}$$

where $\alpha > 0, \beta > 0$ are the shape and scale parameter respectively, where Z has a Logarithmic distribution which has the following probability mass function:

$$p(Z = z) = \frac{(1-p)^z}{-z \ln p}, \quad 0 < p < 1, z = 1, 2, \cdots. \tag{2}$$

Suppose that the variables Y_i and Z are independent, then we have the probability density function of $X|Z = z$,

$$f(x|z) = \alpha\beta z x^{\alpha-1} e^{-\beta z x^{\alpha}}, \quad x > 0,$$

then X follows Weibull-Logarithmic (WL) distribution and its probability density function is given as follows

$$f(x) = \frac{\alpha\beta(p-1)x^{\alpha-1}}{(p + e^{\beta x^{\alpha}} - 1)\ln p}, \quad x > 0. \tag{3}$$

The cumulative distribution function of $WL(\alpha, \beta, p)$ is given by

$$F(x) = 1 - \frac{\ln\left(1 - (1-p)e^{-\beta x^{\alpha}}\right)}{\ln p}, \quad x > 0. \tag{4}$$

2.2 Stochastic ordering

Stochastic order is basic concept in statistics and it measures the notion of one random variable being "bigger" than the other. If $F_X(x) \geq F_Y(x)$ for all real x, it is said that X is less than Y in the ususal stochastic order ($X \prec_{st} Y$). If $h_X(x) \geq h_Y(x)$ for all $x \geq 0$, where $h(x) = f(x)/(1 - F(x))$ is the hazard rate function, it is said that X is less than Y in the hazard rate order ($X \prec_{hr} Y$). If $f_X(x)/f_Y(x)$ increases in x over the union of the supports of X and Y, it is said that X is less than Y in the likelihood ratio order ($X \prec_{lr} Y$). It is clear that $X \prec_{lr} Y \Rightarrow X \prec_{hr} \Rightarrow X \prec_{st} Y$, see [13] for more details.

Theorem 2.1. *Let $X \sim WL(\alpha, \beta, p_1)$ and $Y \sim WL(\alpha, \beta, p_2)$, where $0 < p_2 < p_1 < 1$, then we have $Y \prec_{lr} X$, $Y \prec_{hr} X$ and $Y \prec_{st} X$.*

Proof. Consider the ratio of two densities

$$U(x) = \frac{f_X(x)}{f_Y(x)} = \frac{(p_1 - 1)\left(p_2 + e^{\beta x^\alpha} - 1\right) \ln p_2}{(p_2 - 1)\left(p_1 + e^{\beta x^\alpha} - 1\right) \ln p_1}. \tag{5}$$

Taking the derivative with respect to x,

$$U'(x) = \frac{\alpha\beta\,(p_1 - 1)\,(p_1 - p_2)\,x^{\alpha-1} e^{\beta x^\alpha} \ln p_2}{(p_2 - 1)\left(p_1 + e^{\beta x^\alpha} - 1\right)^2 \ln p_1}. \tag{6}$$

If $p_1 > p_2$, $U'(x) > 0$, then $U(x)$ is increasing at x. The proof is completed. □

2.3 Mean, variance and median

Consider the WL distribution $X \sim WP(\alpha, \beta, p)$, the kth moment of X is as follows, for $k = 1, 2, \ldots$

$$\mu_k = \mathbb{E}(X^k) = k \int_0^\infty x^{k-1} \bar{F}(x) dx = -\frac{\Gamma\left(\frac{k+\alpha}{\alpha}\right) \operatorname{polylog}(\frac{k+\alpha}{\alpha}, (1-p))}{\beta^{\frac{k}{\alpha}} \ln p}, \tag{7}$$

where polylog(.) is a polylogarithm function defined as:

$$\operatorname{polylog}(s, z) = \sum_{k=1}^{\infty} \frac{z^k}{k^s}.$$

Therefore the mean and variance of the WL distribution are

$$E(X) = -\frac{\Gamma\left(\frac{\alpha+1}{\alpha}\right) \operatorname{polylog}(\frac{\alpha+1}{\alpha}, (1-p))}{\beta^{1/\alpha} \ln p}, \tag{8}$$

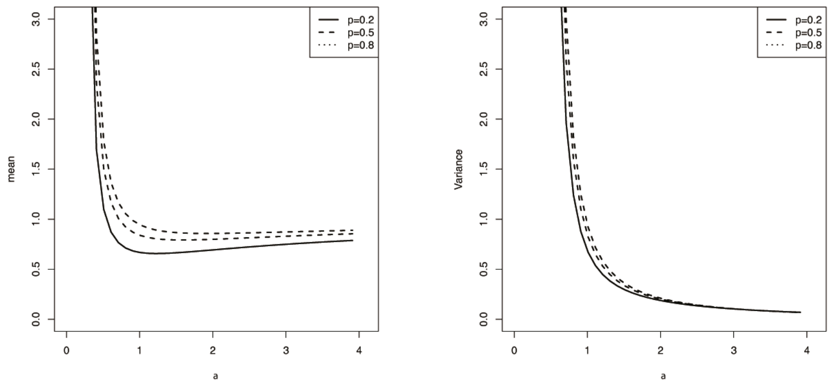
and

$$Var(X) = -\frac{\Gamma\left(1 + \frac{1}{\alpha}\right)^2 \operatorname{polylog}(1 + \frac{1}{\alpha}, (1-p))^2 + \Gamma\left(\frac{\alpha+2}{\alpha}\right) \operatorname{polylog}(\frac{\alpha+2}{\alpha}, (1-p)) \ln p}{\beta^{2/\alpha} (\ln p)^2}. \tag{9}$$

Figures 1a and 1b display the population mean and variance of the $WL(\alpha, \beta = 1, p)$ distribution.

The cdf of the $WL(\alpha, \beta, p)$ distribution is given in (4). The quantile $x_q = F^{-1}(q)$ of the $WL(\alpha, \beta, p)$ distribution is

$$x_q = \left(\frac{1}{\beta} \ln\left(\frac{(1-p)p^q}{p^q - p}\right)\right)^{1/\alpha}.$$



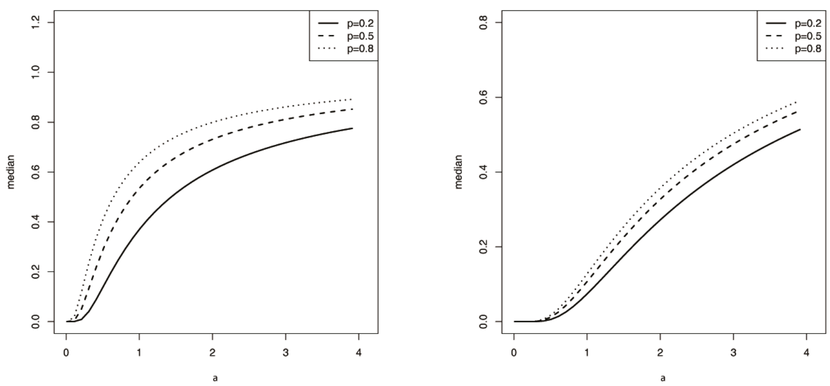
(a) (b)

Figure 1: (a) Mean of the $WL(\alpha, \beta = 1, p)$ distribution; (b) Variance of the $WL(\alpha, \beta = 1, p)$ distribution

In particular, the median ($q = \frac{1}{2}$) of the $WL(\alpha, \beta, p)$ distribution is given by

$$x_m = \left(\frac{\ln\left(\sqrt{p} + 1\right)}{\beta} \right)^{1/\alpha}. \tag{10}$$

Figures 2a and 2b display the population median of the $WL(\alpha, \beta = 1, p)$ and $WL(\alpha, \beta = 5, p)$ distribution.



(a) (b)

Figure 2: (a) Median of the $WL(\alpha, \beta = 1, p)$ distribution; (b) Median of the $WL(\alpha, \beta = 5, p)$ distribution.

§3 Sampling distributions of order statistics

Suppose $X_1, X_2, \ldots, X_n$ form a random sample from the $WL(\alpha, \beta, p)$ distribution. It is known that the sample mean $(X_1 + \cdots + X_n)/n$ converges to the normal distribution as $n \to \infty$ based on the central limit theorem. In this section, we are interested in the asymptotic distributions of the sample minima and maxima, that is, $X_{1:n} = \min(X_1, \ldots, X_n)$ and $X_{n:n} = \max(X_1, \ldots, X_n)$. These two order statistics show the lifetime of series and parallel system and have useful applications in practice.

The probability density function of sample minima $X_{1:n}$ is

$$f_{1:n}(x) = n[\frac{\ln\left(1-(1-p)e^{-\beta x^{\alpha}}\right)}{\ln p}]^{n-1}\frac{\alpha\beta(p-1)x^{\alpha-1}}{(p+e^{\beta x^{\alpha}}-1)\ln p}, \quad x > 0.$$

The probability density function of sample maxima $X_{n:n}$ is

$$f_{n:n}(x) = n[1-\frac{\ln\left(1-(1-p)e^{-\beta x^{\alpha}}\right)}{\ln p}]^{n-1}\frac{\alpha\beta(p-1)x^{\alpha-1}}{(p+e^{\beta x^{\alpha}}-1)\ln p}, \quad x > 0.$$

Theorem 3.1. *Suppose $X_{1:n}$ and $X_{n:n}$ are the smallest and largest order statistics from the $WL(\alpha, \beta, p)$ distribution, then we have*
(1) $\lim_{n\to\infty} P(X_{1:n} \le b_n^* t) = 1 - e^{-t^{\alpha}}, t > 0$, where the constant $b_n^* = [\frac{1}{\beta}\ln(\frac{(1-p)p^{1/n}}{p^{1/n}-p})]^{1/\alpha}$.
(2) $\lim_{n\to\infty} P(X_{n:n} \le a_n + b_n x) = e^{-e^{-x}}, -\infty < x < \infty$, where the constant $a_n = [\frac{1}{\beta}\ln(\frac{(p-1)n}{\ln p})]^{1/\alpha}$ and $b_n = a_n^{1-\alpha}/(\alpha\beta)$.

Proof. (1) We apply the asymptotic results for $X_{1:n}$ as follows (See [2]): For the sample minima $X_{1:n}$, we have, for $t > 0, c > 0$,

$$\lim_{n\to\infty} P(X_{1:n} \le a_n^* + b_n^* t) = 1 - e^{-t^{c}},$$

(it is called Weibull type) where $b_n^* = F^{-1}(1/n) - F^{-1}(0)$ and $a_n^* = F^{-1}(0)$ iff $F^{-1}(0) < \infty$ and

$$\lim_{\epsilon\to 0^+} \frac{F(F^{-1}(0)+\epsilon t)}{F(F^{-1}(0)+\epsilon)} = t^c.$$

As for the $WL(\alpha, \beta, p)$ distribution, its cdf is

$$F(x) = 1 - \frac{\ln\left(1-(1-p)e^{-\beta x^{\alpha}}\right)}{\ln p}, \quad x > 0.$$

Let $F(x) = 0$, we get $x = 0$. Thus $F^{-1}(0) = 0$ is finite. Furthermore,

$$\lim_{\epsilon\to 0^+} \frac{F(0+\epsilon t)}{F(0+\epsilon)} = t \lim_{\epsilon\to 0^+} \frac{f(\epsilon t)}{f(\epsilon)} = t^{\alpha}.$$

Thus, we have $c = \alpha$, $a_n^* = 0$ and $b_n^* = F^{-1}(1/n) = [\frac{1}{\beta}\ln(\frac{(1-p)p^{1/n}}{p^{1/n}-p})]^{1/\alpha}$ which is the $\frac{1}{n}$th quantile.

(2) For the sample maxima $X_{n:n}$,

$$P(X_{n:n} \le a_n + b_n x) = F^n(a_n + b_n x) = [1 - (1 - F(a_n + b_n x))]^n \simeq \exp\{-n(1 - F(a_n + b_n x))\}$$

where $-\infty < x < \infty$. It is sufficient to prove that

$$n(1 - F(a_n + b_n x)) \to e^{-x}$$

as $n \to \infty$. In fact,

$$\begin{aligned} n(1-F(a_n+b_nx)) &= n\frac{\ln\left[1-(1-p)e^{-\beta(a_n+b_nx)^{\alpha}}\right]}{\ln p} \\ &\simeq n\frac{-(1-p)e^{-\beta(a_n+b_nx)^{\alpha}}}{\ln p} \\ &\simeq n\frac{-(1-p)e^{-\beta a_n^{\alpha}(1+\alpha\frac{b_n}{a_n}x)}}{\ln p} \\ &= e^{-x} \end{aligned}$$

as $n \to \infty$. This completes the proof. □

§4 Estimation and inference

4.1 Maximum likelihood estimation

In this section, we discuss the maximum likelihood estimation about the parameters (α, β, p) of the WL distribution. Suppose $y_{obs} = \{x_1, x_2, \ldots, x_n\}$ forms a random sample from the $WL(\alpha, \beta, p)$ distribution, then the log-likelihood function is

$$\begin{aligned} l &= \log\prod_{i=1}^{n} f_X(x_i) \\ &= -\sum_{i=1}^{n}\log\left(e^{\beta x_i^{\alpha}}+p-1\right)+(\alpha-1)\sum_{i=1}^{n}\log(x_i) \\ &\quad +n(\log(\alpha)+\log(\beta)+\log(1-p)-\log(-\log(p))). \end{aligned} \tag{11}$$

The scores functions are

$$\frac{\partial l}{\partial \alpha} = -\sum_{i=1}^{n}\frac{\beta x_i^{\alpha}\log(x_i)\,e^{\beta x_i^{\alpha}}}{e^{\beta x_i^{\alpha}}+p-1}+\sum_{i=1}^{n}\log(x_i)+\frac{n}{\alpha} \tag{12}$$

$$\frac{\partial l}{\partial \beta} = \frac{n}{\beta}-\sum_{i=1}^{n}x_i^{\alpha}-\sum_{i=1}^{n}\frac{(1-p)x_i^{\alpha}}{e^{\beta x_i^{\alpha}}+p-1} \tag{13}$$

$$\frac{\partial l}{\partial p} = n\left(\frac{1}{p-1}-\frac{1}{p\log(p)}\right)-\sum_{i=1}^{n}\frac{1}{e^{\beta x_i^{\alpha}}+p-1}. \tag{14}$$

Maximizing the likelihood function and the estimates of the parameters are obtained. Let the scores functions be zeros and solve the equations. However, there do not exist analytical roots. Therefore, the estimates can only be found by some numerical methods such as Newton-Raphson procedure. [6] fixed the parameter $\alpha = \alpha_0$ and provided two conditions for the solutions. In the following, we present some more general theoretical results.

Theorem 4.1. *Suppose $l_i(\alpha, \beta, p, y_{obs}), i \in \{1, 2, 3\}$ denote the right hand side of equations (12)–(14), respectively, and let $\tilde{x} = (1/n)\sum_{i=1}^{n} x_i^{\alpha}$, then the following properties hold:*
(1) If β and p are given, then the root of $l_1(\alpha, \beta, p, y_{obs}) = 0$, $\hat{\alpha}$, lies in the interval $(0, \infty)$ and is unique.
(2) If α and p are given, then the root of $l_2(\alpha, \beta, p, y_{obs}) = 0$, $\hat{\beta}$, lies in the interval $(p\tilde{x}^{-1}, \tilde{x}^{-1})$

and is unique.
(3) If α and β are known, then for $\sum_{i=1}^{n} e^{-\beta x_i^\alpha} > n/2$, the equation $l_3(\alpha, \beta, p, y_{obs}) = 0$ has at least one root.

Proof. (1) Noting that $\lim_{\alpha\to 0} l_1(\alpha, \beta, p, y_{obs}) = \infty$, $\lim_{\alpha\to\infty} l_1(\alpha, \beta, p, y_{obs}) = -\infty$ as $x_i \geq 1$ for at least one $i \in \{1, 2, \ldots, n\}$, $\lim_{\alpha\to\infty} l_1(\alpha, \beta, p, y_{obs}) = \sum_{i=1}^{n} \log(x_i) < 0$ as $0 < x_i < 1$ for all i. Thus, there exists at least one root of $l_1(\alpha, \beta, p, y_{obs}) = 0$ in the interval $(0, \infty)$. It can be shown that

$$\frac{\partial l_1(\alpha, \beta, p, y_{obs})}{\partial \alpha} = \sum_{i=1}^{n} \frac{\beta x_i^\alpha \log^2(x_i)\, e^{\beta x_i^\alpha} \left[\beta(1-p)x_i^\alpha + 1 - p - e^{\beta x_i^\alpha}\right]}{\left(e^{\beta x_i^\alpha} + p - 1\right)^2} - \frac{n}{\alpha^2}.$$

Since $\beta(1-p)x_i^\alpha + (1-p) - e^{\beta x_i^\alpha} < \beta x_i^\alpha + 1 - e^{\beta x_i^\alpha} < 0$, thus, $l_1(\alpha, \beta, p, y_{obs})$ is decreasing in α, the root is unique.

(2) Let $w(\beta) = \sum_{i=1}^{n} \frac{(1-p)x_i^\alpha}{e^{\beta x_i^\alpha} + p - 1}$. It is obvious that $w(\beta)$ is a strictly decreasing function and $\lim_{\beta\to\infty} w(\beta) = 0$. Then,

$$l_2(\alpha, \beta, p, y_{obs}) < \frac{n}{\beta} - \sum_{i=1}^{n} x_i^\alpha$$

and hence $l_2(\alpha, \beta, p, y_{obs}) < 0$ when $\beta > \tilde{x}^{-1}$.

Furthermore, we have $\lim_{\beta\to 0} w(\beta) = (1-p)/p \sum_{i=1}^{n} x_i^\alpha$. Then,

$$l_2(\alpha, \beta, p, y_{obs}) > \frac{n}{\beta} - \sum_{i=1}^{n} x_i^\alpha - \frac{1-p}{p} \sum_{i=1}^{n} x_i^\alpha = \frac{n}{\beta} - \frac{1}{p} \sum_{i=1}^{n} x_i^\alpha$$

and hence $l_2(\alpha, \beta, p, y_{obs}) > 0$ when $\beta < p\tilde{x}^{-1}$. Therefore, there exists at least one root of $l_2(\alpha, \beta, p, y_{obs}) = 0$ in the interval $(p\tilde{x}^{-1}, \tilde{x}^{-1})$.

As for the uniqueness of the root, we consider

$$\frac{\partial l_2(\alpha, \beta, p, y_{obs})}{\partial \beta} = \sum_{i=1}^{n} \frac{(1-p)x_i^{2\alpha} e^{\beta x_i^\alpha}}{\left(e^{\beta x_i^\alpha} + p - 1\right)^2} - \frac{n}{\beta^2}.$$

For all $\beta > 0$, the function $\partial l_2(\alpha, \beta, p, y_{obs})/\partial\beta$ is not always monotonic. However, if β^* is a root of $\partial l_2(\alpha, \beta, p, y_{obs})/\partial\beta = 0$, then

$$\sum_{i=1}^{n} \frac{(1-p)\beta^* x_i^{2\alpha} e^{\beta^* x_i^\alpha}}{\left(e^{\beta^* x_i^\alpha} + p - 1\right)^2} = \frac{n}{\beta^*}$$

and

$$l_2(\alpha, \beta^*, p, y_{obs}) = \sum_{i=1}^{n} \frac{x_i^\alpha e^{\beta^* x_i^\alpha} \left[\beta^*(1-p)x_i^\alpha + (1-p) - e^{\beta^* x_i^\alpha}\right]}{\left(e^{\beta^* x_i^\alpha} + p - 1\right)^2}.$$

Since $\beta^*(1-p)x_i^\alpha + (1-p) - e^{\beta^* x_i^\alpha} < \beta^* x_i^\alpha + 1 - e^{\beta^* x_i^\alpha} < 0$, so for all β^*, $l_2(\alpha, \beta^*, p, y_{obs}) < 0$ at its stationary points. Taking into account that $\lim_{\beta^*\to\infty} l_2(\alpha, \beta^*, p, y_{obs}) = -\sum_{i=1}^{n} x_i^\alpha < 0$, the uniqueness of the root is proven.

(3) It is clear that $\lim_{p\to 0} l_3(\alpha, \beta, p, y_{obs}) = \infty$ and $\lim_{p\to 1} l_3(\alpha, \beta, p, y_{obs}) = n/2 - \sum_{i=1}^{n} e^{-\beta x_i^\alpha} < 0$, thus, there exists at least one solution of $l_3(\alpha, \beta, p, y_{obs}) = 0$ for $0 < p < 1$. □

4.2 An Expectation–Maximization (EM) algorithm

In the case of the model involves unobserved latent variables, an Expectation–Maximization (EM) algorithm [8]) is a powerful tool to obtain maximum likelihood estimates of parameters.

The EM algorithm consists of two steps, an expectation (E) step and a maximization (M) step. The E-step produces a function for the expectation of the log-likelihood and the M-step calculates the parameters which maximize the function obtained on the previous E-step. The estimates are then adopted to obtain the distribution of the latent variables in the next E-step.

Let X, Z denote the observed and the missing data. The density function of (X, Z) is

$$f(x,z) = p(z)f(x|z) = -\frac{(1-p)^z}{\ln p}\alpha\beta x^{\alpha-1}e^{-\beta z x^{\alpha}}, \quad z = 1, 2, \ldots, x > 0.$$

It is trivial to compute the conditional expectation of $(Z|X)$ using the pdf

$$p(z|x) = (1-p)^{z-1}\left(p + e^{\beta x^{\alpha}} - 1\right)e^{-\beta z x^{\alpha}}, \quad z = 1, 2, \ldots,$$

that is,

$$\mathbb{E}(Z|X) = \frac{1}{1-(1-p)e^{-\beta x^{\alpha}}}.$$

The iteration is completed with the M-step which is essentially-full data maximum likelihood, with the $Z's$ replaced by $\mathbb{E}(Z|X)$. Threfore, the iteration is as follows:

$$\alpha^{(t+1)} = n\left[\sum_{i=1}^{n}\log(x_i)\left(\frac{\beta x_i^{\alpha^{(t+1)}}}{1-(1-p^{(t)})e^{-\beta^{(t)}x_i^{\alpha^{(t)}}}} - 1\right)\right]^{-1},$$

$$\beta^{(t+1)} = n\left[\sum_{i=1}^{n}\frac{x_i^{\alpha^{(t+1)}}}{1-(1-p^{(t)})e^{-\beta^{(t)}x_i^{\alpha^{(t)}}}}\right]^{-1},$$

$$p^{(t+1)} = \frac{n(p^{(t+1)}-1)}{\log(p^{(t+1)})\sum_{i=1}^{n}\{1-(1-p^{(t)})e^{-\beta^{(t)}x_i^{\alpha^{(t)}}}\}^{-1}}.$$

4.3 Asymptotic properties of MLEs

By the classic statistics theory, under some regular conditions, the distribution of the MLE converges to the multivariate normal distribution which has the mean (α, β, p) and its covariance equals to the inverse of the Fisher matrix, see [7] for details. We can use the normal distribution to construct approximate confidence intervals for α, β and p.

Let $I = I(\alpha, \beta, p; y_{obs})$ be the observed information matrix. The elements $I_{ij}, i, j = 1, 2, 3$ are given as:

$$I_{11} = \sum_{i=1}^{n}\frac{\beta^2 x_i^{2\alpha}\log^2(x_i)\,e^{\beta x_i^{\alpha}}}{e^{\beta x_i^{\alpha}}+p-1} - \sum_{i=1}^{n}\frac{\beta^2 x_i^{2\alpha}\log^2(x_i)\,e^{2\beta x_i^{\alpha}}}{\left(e^{\beta x_i^{\alpha}}+p-1\right)^2} + \sum_{i=1}^{n}\frac{\beta x_i^{\alpha}\log^2(x_i)\,e^{\beta x_i^{\alpha}}}{e^{\beta x_i^{\alpha}}+p-1} + \frac{n}{\alpha^2},$$

$$I_{12} = I_{21} = \sum_{i=1}^{n}\frac{x_i^{\alpha}\log(x_i)\,e^{\beta x_i^{\alpha}}}{e^{\beta x_i^{\alpha}}+p-1} + \sum_{i=1}^{n}\frac{\beta x_i^{2\alpha}\log(x_i)\,e^{\beta x_i^{\alpha}}}{e^{\beta x_i^{\alpha}}+p-1} - \sum_{i=1}^{n}\frac{\beta x_i^{2\alpha}\log(x_i)\,e^{2\beta x_i^{\alpha}}}{\left(e^{\beta x_i^{\alpha}}+p-1\right)^2},$$

$$I_{13} = I_{31} = -\sum_{i=1}^{n} \frac{\beta x_i^\alpha \log(x_i) e^{\beta x_i^\alpha}}{\left(e^{\beta x_i^\alpha} + p - 1\right)^2},$$

$$I_{22} = \sum_{i=1}^{n} \frac{x_i^{2\alpha} e^{\beta x_i^\alpha}}{e^{\beta x_i^\alpha} + p - 1} - \sum_{i=1}^{n} \frac{x_i^{2\alpha} e^{2\beta x_i^\alpha}}{\left(e^{\beta x_i^\alpha} + p - 1\right)^2} + \frac{n}{\beta^2},$$

$$I_{23} = I_{32} = -\sum_{i=1}^{n} \frac{x_i^\alpha e^{\beta x_i^\alpha}}{\left(e^{\beta x_i^\alpha} + p - 1\right)^2},$$

$$I_{33} = -\sum_{i=1}^{n} \frac{1}{\left(e^{\beta x_i^\alpha} + p - 1\right)^2} - n\left(\frac{1}{p^2 \log^2(p)} + \frac{1}{p^2 \log(p)} - \frac{1}{(p-1)^2}\right).$$

Taking the expectation $J = \mathbb{E}[I(\alpha, \beta, p; y_{obs})]$ with respect to the distribution of X, we obtain the fisher information matrix:

$$J(\theta, p) = n \begin{pmatrix} J_{11} & J_{12} & J_{13} \\ J_{21} & J_{22} & J_{23} \\ J_{31} & J_{32} & J_{33} \end{pmatrix}$$

where

$$J_{11} = \mathbb{E}\frac{\beta^2 X^{2\alpha} \log^2(X) e^{\beta X^\alpha}}{e^{\beta X^\alpha} + p - 1} - \mathbb{E}\frac{\beta^2 X^{2\alpha} \log^2(X) e^{2\beta X^\alpha}}{(e^{\beta X^\alpha} + p - 1)^2} + \mathbb{E}\frac{\beta X^\alpha \log^2(X) e^{\beta X^\alpha}}{e^{\beta X^\alpha} + p - 1} + \frac{1}{\alpha^2},$$

$$J_{12} = J_{21} = \mathbb{E}\frac{X^\alpha \log(X) e^{\beta X^\alpha}}{e^{\beta X^\alpha} + p - 1} + \mathbb{E}\frac{\beta X^{2\alpha} \log(X) e^{\beta X^\alpha}}{e^{\beta X^\alpha} + p - 1} - \mathbb{E}\frac{\beta X^{2\alpha} \log(X) e^{2\beta X^\alpha}}{(e^{\beta X^\alpha} + p - 1)^2},$$

$$J_{13} = J_{31} = -\mathbb{E}\frac{\beta X^\alpha \log(X) e^{\beta X^\alpha}}{(e^{\beta X^\alpha} + p - 1)^2},$$

$$J_{22} = \mathbb{E}\frac{X^{2\alpha} e^{\beta X^\alpha}}{e^{\beta X^\alpha} + p - 1} - \mathbb{E}\frac{X^{2\alpha} e^{2\beta X^\alpha}}{(e^{\beta X^\alpha} + p - 1)^2} + \frac{1}{\beta^2},$$

$$J_{23} = J_{32} = -\mathbb{E}\frac{X^\alpha e^{\beta X^\alpha}}{(e^{\beta X^\alpha} + p - 1)^2},$$

$$J_{33} = -\mathbb{E}\frac{1}{(e^{\beta X^\alpha} + p - 1)^2} - \frac{1}{p^2 \log^2(p)} - \frac{1}{p^2 \log(p)} + \frac{1}{(p-1)^2}.$$

The asymptotic variance–covariance matrix of the MLEs can be obtained from the inverse of $J(\alpha, \beta, p)$, evaluated at $\hat{\alpha}$, $\hat{\beta}$ and $\hat{p}$. Note that the inverse of the observed information matrix is a consistent estimator of J^{-1}, thus another estimates can be calculated from it.

4.4 Simulation studies

In this section, we conduct a simulation study to assess the performance of the approximation of the covariances and variances of the MLEs calculated from the information matrix.

For each value of (α, β, p), we compute the parameter estimates by the EM algorithm in Section 4.2 with various initial values. The iteration is stopped when the absolute differences between successive estimates are smaller than 10^{-5}.

Table 1 shows the simulated values of $Var(\hat{\alpha})$, $Var(\hat{\beta})$, $Var(\hat{p})$, $Cov(\hat{\alpha},\hat{\beta})$, $Cov(\hat{\alpha},\hat{p})$ and $Cov(\hat{\beta},\hat{p})$ as well as the approximate values computed by averaging the values obtained from the observed and expected information matrices.

It is clear that for large numbers of sample size n, the approximate values computed from the expected and observed information matrices are almost equal to the corresponding simulated ones. The approximation gets quite accurate as sample size n is larger. For large numbers of sample size n, the covariances and variances of the MLEs computed from the observed information matrix is close to that of the expected information matrix as expected.

We also conduct simulations to study the convergence of the proposed EM algorithm in Section 4.2. For each of the three values of the parameters, 1000 samples of size 100 and 500 sampled from the WL distribution are generated.

Table 2 provides the averages of the 1000 MLEs, $av(\hat{\alpha}), av(\hat{\beta}), av(\hat{p})$ and their corresponding standard errors. From Table 2, we can see that the convergence is satisfied in all cases, even when the initial values are bad. The results verify the stability of the EM algorithm. The EM estimates performed consistently. As expected, when n is larger, the standard errors of the MLEs decrease.

§5 Applications

In this section, we apply our model to a real dataset, the stress-rupture data set. The data set was previously studied by [3], [5], [10] and [12]. It consists of the life of fatigue fracture of Kevlar 49/epoxy that are subject to the pressure at the 90% level.

We fit the $WL(\alpha,\beta,p)$, $EL(\beta,p)$, $WG(\alpha,\beta,p)$ and $Weibull(\alpha,\beta)$ distributions to the dataset. The performances of the distributions are discussed.

The mentioned probability density functions are given as follows:

$$EL: \quad f(x|\Theta_1) = \frac{1}{-\log(p)} \frac{\beta(1-p)e^{-\beta x}}{1-(1-p)e^{-\beta x}}, \quad \Theta_1 = (\beta, p), \quad x > 0,$$

$$WG: \quad f(x|\Theta_2) = \frac{\alpha\beta^\alpha(1-p)x^{\alpha-1}e^{-(\beta x)^\alpha}}{[1-pe^{-(\beta x)^\alpha}]^2}, \quad \Theta_2 = (\alpha,\beta,p), \quad x > 0,$$

$$Weibull: \quad f(x|\Theta_3) = \alpha\beta x^{\alpha-1}e^{-\beta x^\alpha}, \quad \Theta_3 = (\alpha,\beta), \quad x > 0.$$

We obtain the MLEs of the parameters and the results are shown in Table 3. We use the Akaike information criterion (AIC) to assess the goodness of fit of the distributions. Given a class of candidate distributions, the preferred model is the one with the smallest AIC value.

We also present the Kolmogorov-Smirnov (K-S) statistics and the corresponding p-values for these distributions. Some other statistics such as the Cramer-Von Mises and Anderson-Darling statistics ([9]) are also obtained.

From Table 3, for the fracture data set, AIC displays that WL model is a best fit. It has the smallest AIC and the highest likelihood values. The K-S statistic and the other two statistics also take the smallest values under the WL model. Figure 3 displays the probability-probability (P-P) plot for the dataset.

Table 1: The simulated values , values obtained from expected information , values obtained from observed information❿ of covariances and variances of the MLEs

n	(α, β, p)	$Var(\hat{\alpha})$	$Cov(\hat{\alpha}, \hat{\beta})$	$Cov(\hat{\alpha}, \hat{p})$	$Var(\hat{\beta})$	$Cov(\hat{\beta}, \hat{p})$	$Var(\hat{p})$
50		0.0047	-0.0096	-0.0090	0.0234	0.0213	0.0227
	(1, 1, 0.2)	0.0034	-0.0064	-0.0052	0.0167	0.0134	0.0143
		0.0063	-0.0132	-0.0111	0.0333	0.0290	0.0321 ❿
		0.0034	-0.0103	-0.0101	0.0452	0.0377	0.0403
	(1, 2, 0.4)	0.0029	-0.0075	-0.0080	0.0347	0.0302	0.0359
		0.0058	-0.0178	-0.0182	0.0718	0.0695	0.0827
		0.0124	-0.0114	-0.0104	0.0138	0.0115	0.0110
	(2, 1, 0.2)	0.0243	-0.0257	-0.0212	0.0325	0.0267	0.0258
		0.0250	-0.0263	-0.0212	0.0329	0.0267	0.0254
		0.0131	-0.0190	-0.0215	0.0427	0.0408	0.0451
	(2, 2, 0.4)	0.0230	-0.0353	-0.0371	0.0719	0.0719	0.0896
		0.0232	-0.0350	-0.0355	0.0703	0.0672	0.0799
200		0.0027	-0.0050	-0.0040	0.0119	0.0091	0.0079
	(1, 1, 0.2)	0.0017	-0.0032	-0.0024	0.0081	0.0056	0.0047
		0.0032	-0.0065	-0.0048	0.0156	0.0116	0.0100
		0.0021	-0.0065	-0.0057	0.0260	0.0204	0.0178
	(1, 2, 0.4)	0.0015	-0.0038	-0.0038	0.0171	0.0140	0.0146
		0.0029	-0.0089	-0.0088	0.0354	0.0324	0.0343
		0.0072	-0.0078	-0.0055	0.0102	0.0068	0.0050
	(2, 1, 0.2)	0.0122	-0.0128	-0.0100	0.0159	0.0122	0.0104
		0.0120	-0.0126	-0.0098	0.0155	0.0118	0.0101
		0.0072	-0.0108	-0.0097	0.0244	0.0182	0.0159
	(2, 2, 0.4)	0.0118	-0.0178	-0.0172	0.0352	0.0312	0.0320
		0.0122	-0.0182	-0.0174	0.0357	0.0312	0.0315
1000		0.0021	-0.0042	-0.0031	0.0097	0.0069	0.0055
	(1, 1, 0.2)	0.0011	-0.0021	-0.0016	0.0054	0.0039	0.0033
		0.0020	-0.0042	-0.0033	0.0103	0.0078	0.0067
		0.0011	-0.0030	-0.0028	0.0144	0.0109	0.0099
	(1, 2, 0.4)	0.0010	-0.0025	-0.0026	0.0116	0.0096	0.0098
		0.0019	-0.0060	-0.0060	0.0239	0.0218	0.0223
		0.0062	-0.0063	-0.0049	0.0082	0.0059	0.0047
	(2, 1, 0.2)	0.0082	-0.0086	-0.0067	0.0105	0.0080	0.0067
		0.0082	-0.0085	-0.0065	0.0104	0.0078	0.0065
		0.0047	-0.0070	-0.0083	0.0154	0.0149	0.0166
	(2, 2, 0.4)	0.0074	-0.0118	-0.0132	0.0249	0.0253	0.0293
		0.0073	-0.0116	-0.0128	0.0245	0.0246	0.0281

Table 2: The means and standard errors of the EM estimator with initial values $(\alpha^{(0)}, \beta^{(0)}, p^{(0)})$ from 1000 samples

n	(α, β, p)	$(\alpha^{(0)}, \beta^{(0)}, p^{(0)})$	$av(\hat{\alpha})$	$av(\hat{\beta})$	$av(\hat{p})$	$se(\hat{\alpha})$	$se(\hat{\beta})$	$se(\hat{p})$
	(1, 1, 0.2)	(1, 1, 0.2)	1.0020	0.9992	0.2208	0.0610	0.1331	0.1178
100	(1, 2, 0.4)	(1, 2, 0.4)	1.0109	1.9881	0.4351	0.0679	0.2432	0.2301
	(2, 2, 0.4)	(2, 2, 0.4)	2.0194	1.9559	0.4539	0.1474	0.2332	0.2346
	(1, 1, 0.2)	(1.5, 1.5, 0.5)	1.0018	1.0207	0.2414	0.0769	0.1758	0.1520
100	(1, 2, 0.4)	(1.5, 2.5, 0.5)	1.0019	2.0499	0.4904	0.0697	0.2481	0.2513
	(2, 2, 0.4)	(2.5, 2.5, 0.5)	2.0076	2.0221	0.4567	0.1310	0.2384	0.2251
	(1, 1, 0.2)	(1, 1, 0.2)	1.0008	0.9998	0.2136	0.0486	0.1087	0.0801
500	(1, 2, 0.4)	(1, 2, 0.4)	1.0014	2.0152	0.4303	0.0510	0.1812	0.1818
	(2, 2, 0.4)	(2, 2, 0.4)	2.0019	1.9838	0.4034	0.0831	0.1554	0.1386
	(1, 1, 0.2)	(1.5, 1.5, 0.5)	0.9975	1.0269	0.2364	0.0547	0.1262	0.1040
500	(1, 2, 0.4)	(1.5, 2.5, 0.5)	1.0076	1.9887	0.4114	0.0510	0.1737	0.1597
	(2, 2, 0.4)	(2.5, 2.5, 0.5)	2.0083	2.0117	0.4291	0.1004	0.1671	0.1534

Table 3: MLEs (with (SE)) of the WL, EL, WG and Weibull distributions for the fracture dataset

Model	Estimates			log-lik	AIC	K-S stat.	p-value	Cramer	Darling
WL	0.9267	1.0064	0.7901	−102.5598	211.1196	0.0891	0.8175	0.1958	1.1821
	(0.1091)	(0.2969)	(0.3083)						
EL	0.9255	0.8100		−103.9421	211.8842	0.0990	0.7052	0.1960	1.2922
	(0.2034)	(0.6218)							
WG	0.9276	1.0037	0.2138	−102.6942	211.3884	0.0891	0.8175	0.1963	1.1845
	(0.1230)	(0.3829)	(0.5686)						
Weibull	0.9767	0.8823		−103.7927	211.5854	0.1089	0.5870	0.1977	1.3884
	(0.0755)	(0.0908)							

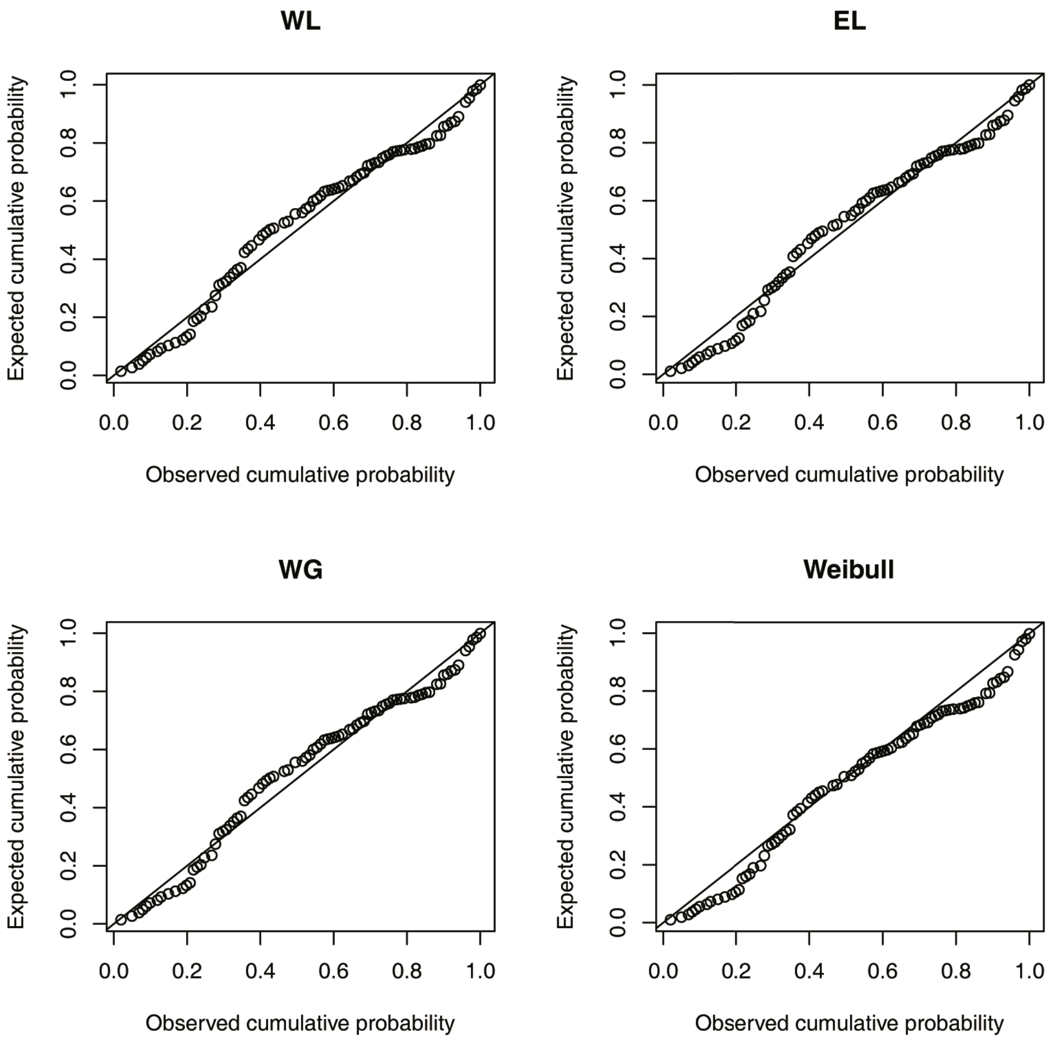


Figure 3: P-P plots for the fracture dataset

References

[1] K Adamidis, S Loukas. *A lifetime distribution with decreasing failure rate*, Statist Probab Lett, 1998, 39(1): 35-42.

[2] B Arnold, N Balakrishnan, H Nagaraja. *A First Course in Order Statistics*, Classics in Applied Mathematics, vol 54, SIAM, 1992.

[3] R Barlow, R Toland, T Freeman. *A Bayesian analysis of the stress-rupture life of kevlar/epoxy spherical pressure vessels*, In: *Accelerated Life Testing and Experts' Opinions in Reliability*, C A Clarotti, D V Lindley, eds, 1988.

[4] W Barreto-Souza, A L de Morais, G M Cordeiro. *The Weibull-geometric distribution*, J Stat Comput Simul, 2011, 81(5): 645-657.

[5] L Castro, H Gómez, M Valenzuela. *Epsilon half-normal model: properties and inference*, Comput Statist Data Anal, 2012.

[6] R Ciumara, V Preda. *The Weibull-logarithmic distribution in lifetime analysis and its properties*, In: *The XIII International Conference Applied Stochastic Models and Data Analysis*, L Sakalauskas, C Skiadas, and E K Zavadskas, eds, 2009, pp 395-399.

[7] D Cox, D Hinkley. *Theoretical Statistics*, Chapman & Hall/CRC, 1979.

[8] A P Dempster, N M Laird, D B Rubin. *Maximum likelihood from incomplete data via the EM algorithm*, J R Stat Soc Ser B Methodol, 1977, 39(1): 1-38.

[9] D L Evans, J H Drew, L M Leemis. *The distribution of the Kolmogorov–Smirnov, Cramer–Von mises, and Anderson–Darling test statistics for Exponential populations with estimated parameters*, Comm Statist Simulation Comput, 2008, 37(7): 1396-1421.

[10] W Gui, P-H Chen, H Wu. *An epsilon half normal slash distribution and its applications to non-negative measurements*, Open J Optim, 2013, 2(1): 1-8.

[11] N L Johnson, A W Kemp, S Kotz. *Univariate Discrete Distributions*, Vol 444, John Wiley & Sons, 2005.

[12] N Olmos, H Varela, H Gómez, H Bolfarine. *An extension of the half-normal distribution*, Statist Papers, 2012, 53(4): 875-886.

[13] C Ramesh, S Kirmani. *On order relations between reliability measures*, Stoch Models, 1987, 3(1): 149-156.

[14] R Tahmasbi, S Rezaei. *A two-parameter lifetime distribution with decreasing failure rate*, Comput Statist Data Anal, 2008, 52(8): 3889-3901.

[1] Department of Mathematics, Beijing Jiaotong University, Beijing 100044, China.
Email: guiwenhao@gmail.com

[2] Department of Mathematics and Physics, Beijing Institute of Petro-chemical Technology, Beijing 102617, China. Email: zhanghuainian@bipt.edu.cn