Thailand Statistician
January 2017; 15(1): 55-68
http://statassoc.or.th
Contributed paper

Discriminating between the Weibull and Power Lindley Distributions Utilizing Fisher Information

Winai Bodhisuwan *[a] and Yuwadee Sangsanit [a]
[a] Department of Statistics, Kasetsart University, Bangkok 10900, Thailand.
*Corresponding author; e-mail: fsciwnb@ku.ac.th

Received: 7 September 2015
Accepted: 10 March 2016

Abstract

Recently, the power Lindley distribution was proposed by Ghitany et al. (2013). It can be considered as an alternative to the Weibull distribution. In some cases, the Weibull and PL distributions provide similar data fit. In this paper, we discriminate between the Weibull and PL distributions by taking advantage of the Fisher information procedure proposed by Gupta and Kundu (2006). In addition, the Fisher information matrix is used to compute the asymptotic variance of the p^{th} percentile estimators including the asymptotic variances of the median and the average asymptotic variance of the percentile estimators. Furthermore, both of the asymptotic variances are useful as they can provide the further information of the Q-Q plot, especially the variation. In cases of the Weibull and PL distributions, the trends of the asymptotic variance of the *p*th percentile estimators are demonstrated with respect to the shape and the scale parameters. Moreover, we also analyze two datasets and compare the result between the likelihood ratio and the Fisher information procedures.

Keywords: Fisher information matrix, percentile estimators, Weibull distribution, power Lindley distribution.

1. Introduction

Many researchers pay attention to discriminate among any closeness distributions (Alshunnara et al. 2010; Gupta and Kundu 2003, 2006; Kundu and Raqab 2007; Pakyari 2010; Raqaba 2013). Some sets of distributions are considered similarly according to the shape of density function, the behavior of hazard function, and the tail probability. In literature, there are many distributions involved in discrimination problem such as the Weibull and generalized exponential distributions (Gupta and Kundu 2003, 2006) the generalized Rayleigh and log-normal distributions (Kundu and Raqab 2007; Alshunnara et al. 2010) and the generalized Rayleigh and Weibull distributions (Raqaba 2013). It is important to mention that those distributions overlap with each other in some ranges of parameters. Moreover, the fitted distributions for given dataset are quite similar. Therefore, the discrimination problem is necessary to be studied.

Ghitany et al. (2013) proposed the power Lindley (PL) distribution. It is a two component mixture of the Weibull distribution and a generalized gamma distribution. The PL distribution has two parameters which behave as shape and scale. Indeed, the PL distribution has more shape of probability density function than the Weibull distribution, which can broaden the range of application. Its hazard function can be decreasing, increasing and decreasing-increasing-decreasing (Ghitany et al. 2013). Furthermore, the decreasing-increasing-decreasing behavior of hazard function can explain new kind of phenomenon in practice. In application study, Ghitany et al. (2013) fitted the PL distribution to real dataset and compared the result with the Gompert, gamma and the Weibull distributions. Remarkably, the PL distribution provided similar maximum log-likelihood values to the Weibull distribution. Indeed, the Weibull distribution has gained a lot of attention from many researchers (Rinne 1986). There are many books and works discuss the Weibull distribution, including properties, applications, generalization, and competitive distributions. The PL distribution is one of its competitive distributions. In practice, the nearest fitted results of the Weibull and PL distributions may happen such as in reliability analysis. Consequently, we are interested in discriminating between the PL and Weibull distributions.

The Fisher information matrix has various advantages. It can measure the information of the parameter containing in data. In inferential statistics, the Fisher information matrix is associated with the asymptotic properties of the maximum likelihood estimators (Hofmann 2004). Additionally, the factorization of the hazard function can be characterized by the property of the Fisher information (Hofmann et al. 2005). In 2006, Gupta and Kundu applied the Fisher information matrix to discriminate between the Weibull and generalized exponential (GE) distributions (Gupta and Kundu 2006). They found that both Weibull and GE distributions can provide similar fitted result of some particular datasets, but their Fisher information matrices can be quite different. In fact, they use the Fisher information matrix to compute the asymptotic variance of the *p*th percentile estimators including the asymptotic variances of the median and average asymptotic variance of the percentile estimators. Moreover, the p^{th} percentile estimators are functions of MLE, asymptotic variance of the *p*th percentile estimators and asymptotic variance of median provides a framework of discriminating distributions through specific function of MLE. This framework would be useful in practice.

In this paper, above asymptotic variance of the p^{th} percentile is applied to discriminate between the Weibull and PL distributions. The rest of this paper is organized as follows. Section 2 provides fundamental properties of the Weibull and PL distributions. In Section 3, the Fisher information matrix of the Weibull and PL distributions are presented. Furthermore, we give the discrimination procedures considered in this study in Section 4. Section 5 illustrates the trends of the asymptotic variances of the median and average asymptotic variance of the percentile estimators with respect to the parameters of an underlying distribution. Moreover, the application to real data is demonstrated in Section 6.

2. Preliminaries

In this section, we provide fundamental properties of the Weibull and PL distributions. In addition, both of the distributions have two parameters that are shape and scale parameters. The closeness of the Weibull and PL distributions is presented in Figure 1.

2.1. Weibull distribution

The distribution function of the Weibull distribution is

$$F_{WB}(x;\beta,\gamma) = 1 - e^{-(\gamma x)^{\beta}}$$

and the corresponding probability density function is

$$f_{WB}(x;\beta,\gamma) = \beta\gamma^{\beta}x^{\beta-1}e^{-(\gamma x)^{\beta}},$$

where $x > 0$, $\beta > 0$ is the shape parameter, and $\gamma > 0$ is the scale parameter (Rinne 1986).

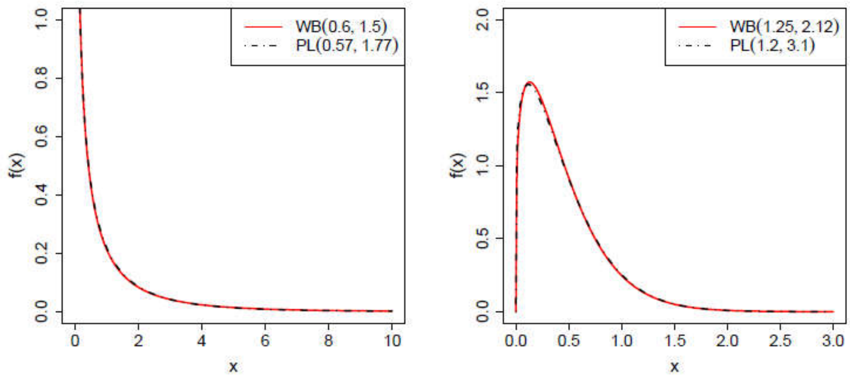

Figure 1 Plots of the Weibull and PL pdfs for some parameter values

The Weibull distribution with shape parameter β and scale parameter γ will be denoted by $WB(\beta,\gamma)$.

The log-likelihood function of the Weibull distribution based on a random sample of size n is

$$\log L_{WB}(\beta,\gamma) = n\log(\beta) + n\beta\log(\gamma) + (\beta-1)\sum_{i=1}^{n}\log(x_i) - \gamma^{\beta}\sum_{i=1}^{n}x_i^{\beta}.$$

The first derivatives of $\log L_{WB}(\beta,\gamma)$ with respect to β and γ are

$$\frac{\partial}{\partial\beta}\log L_{WB}(\beta,\gamma) = \frac{n}{\beta} + n\log(\gamma) + \sum_{i=1}^{n}\log(x_i) - \gamma^{\beta}\sum_{i=1}^{n}x_i^{\beta}\log(\gamma x_i) \tag{1}$$

$$\frac{\partial}{\partial\gamma}\log L_{WB}(\beta,\gamma) = \frac{n\beta}{\gamma} - \beta\gamma^{\beta-1}\sum_{i=1}^{n}x_i^{\beta}. \tag{2}$$

By equating (1) and (2) to zero, we get the nonlinear equations. Consequently, these equations can be solved by a numerical method such as the Newton-Raphson algorithm in optim function of R language (R Core Team 2015) to find the maximum likelihood estimates (MLE) of β and γ.

2.2. Power Lindley distribution

The PL distribution has the following distribution function and probability density function, respectively,

$$F_{PL}(x;\alpha,\lambda) = 1 - \left(1 + \frac{\lambda}{\lambda+1}x^{\alpha}\right)e^{-\lambda x^{\alpha}},$$

and

$$f_{PL}(x;\alpha,\lambda) = \frac{\alpha\lambda^2}{\lambda+1}(1+x^{\alpha})x^{\alpha-1}e^{-\lambda x^{\alpha}},$$

where $x > 0$, $\alpha > 0$ is the shape parameter, and $\lambda > 0$ is the scale parameter. The PL distribution with shape parameter α and scale parameter λ will be written by $PL(\alpha,\lambda)$. The log-likelihood function of the PL distribution based on a random sample of size n is

$$\log L_{PL}(\alpha,\lambda) = n\log(\alpha) + 2n\log(\lambda) - n\log(\lambda+1) + \sum_{i=1}^{n}\log(1+x_i^{\alpha}) - \lambda\sum_{i=1}^{n}x_i^{\alpha} + (\alpha-1)\sum_{i=1}^{n}\log(x_i).$$

By differentiating $\log L_{PL}(\alpha,\lambda)$ with respect to α and λ, respectively, the components of the unit score vector are

$$\frac{\partial}{\partial\alpha}\log L_{PL}(\alpha,\lambda) = \frac{n}{\alpha} + \sum_{i=1}^{n}\frac{x_i^{\alpha}\log(x_i)}{1+x_i^{\alpha}} + \sum_{i=1}^{n}\log(x_i) - \lambda\sum_{i=1}^{n}x_i^{\alpha}\log(x_i) \tag{3}$$

$$\frac{\partial}{\partial\lambda}\log L_{PL}(\alpha,\lambda) = \frac{2n}{\lambda} - \frac{n}{\lambda+1} - \sum_{i=1}^{n}x_i^{\alpha}. \tag{4}$$

Therefore, the MLEs of α and λ can be obtained by equating (3) and (4) to zero and then solving numerically.

3. Fisher information matrix

The information matrix is a crucial part in statistical inference, related to estimation, sufficiency and properties of variance of estimators (Nadarajah 2006). Under certain regularity conditions (Lehmann 1991), the elements of the information matrix are

$$(I)_{jk} = \left(E\left(-\frac{\partial\log f(\boldsymbol{X}:\boldsymbol{\theta})}{\partial\theta_j}\frac{\partial\log f(\boldsymbol{X}:\boldsymbol{\theta})}{\partial\theta_k}\right)\right)$$

where $\boldsymbol{\theta} = (\theta_1,\theta_2,\ldots,\theta_p)$ and $j,k = 1,2,\ldots,p.$

Indeed, the hazard function can be used to obtain the information matrix (Efron and Johnstone 1990).

Let $h(\boldsymbol{X}:\boldsymbol{\theta}) = \dfrac{f(\boldsymbol{X}:\boldsymbol{\theta})}{1-F(\boldsymbol{X}:\boldsymbol{\theta})}$ be the hazard function of $\boldsymbol{X}$. Accordingly, the elements of the information matrix are

$$(I)_{jk} = \left(E\left(-\frac{\partial\log h(\boldsymbol{X}:\boldsymbol{\theta})}{\partial\theta_j}\frac{\partial\log h(\boldsymbol{X}:\boldsymbol{\theta})}{\partial\theta_k}\right)\right).$$

3.1. Fisher information matrix of Weibull

If a random variable X has the Weibull distribution, the Fisher information for the parameter vector $\boldsymbol{\theta}$ is

$$\boldsymbol{I}_{WB}(\boldsymbol{\theta}) = \begin{bmatrix} I_{WB11} & I_{WB12} \\ I_{WB21} & I_{WB22}, \end{bmatrix},$$

where

$$I_{WB11} = \frac{1}{\beta^2}(\psi'(1) + \psi^2(2)),$$

$$I_{WB12} = \frac{1}{\gamma}(1+\psi(1)),$$

$$I_{WB22} = \frac{\beta^2}{\gamma^2}.$$

where $\Gamma(\cdot)$ is the gamma function, $\psi(\cdot) = \Gamma'(\cdot)/\Gamma(\cdot)$ is the digamma function, and $\psi^2(\cdot)$ is the second derivative of the logarithm of the gamma function.

3.2. Fisher information matrix of PL

Let a random variable X have the PL distribution, the Fisher information for the parameter vector $\boldsymbol{\theta}$ is

$$\boldsymbol{I}_{PL}(\boldsymbol{\theta}) = \begin{bmatrix} I_{PL11} & I_{PL12} \\ I_{PL21} & I_{PL21}, \end{bmatrix},$$

where

$$I_{PL11} = \frac{1}{\alpha^2} + \frac{1}{\alpha^2(\beta+1)}\Big\{\beta\big[(\psi(2)-\log(\beta))^2+\zeta(2,2)\big]+2\big[(\psi(3)-\log(\beta))^2+\zeta(2,3)\big]$$
$$-\beta\big[(\psi(1)-\log(\beta))^2+\zeta(2,1)\big]+\beta^2 J(\beta)\Big\},$$

$$I_{PL12} = \frac{\beta\big[\psi(2)-\log(\beta)\big]+2\big[\psi(3)-\log(\beta)\big]}{\alpha\beta(\beta+1)},$$

$$I_{PL22} = \frac{\beta^2+4\beta+2}{\beta^2(\beta+1)^2},$$

where $\zeta(a,b)$ is the Riemann's zeta function written by

$$\zeta(a,b) = \sum_{m=0}^{\infty}\frac{1}{(b+m)^a}, \qquad a>1, b\neq 0,-1,-2,\ldots$$

and

$$J(\beta) = \int_0^{\infty}\frac{(\log t)^2}{1+t}e^{-\beta t}\,dt.$$

4. Discrimination procedures

This section includes two discrimination procedures which are likelihood ratio and Fisher information procedures.

4.1. Likelihood ratio procedure

Suppose a random sample $X_1,\ldots,X_n$ belong to either the Weibull distribution or the PL distribution. Let maximum likelihood of the Weibull distribution and the PL distribution be

$$L_{WB}(\hat{\beta},\hat{\gamma}) = \prod_{i=1}^{n} f_{WB}(x_i;\hat{\beta},\hat{\gamma}),$$

and

$$L_{PL}(\hat{\alpha},\hat{\lambda}) = \prod_{i=1}^{n} f_{PL}(x_i;\hat{\alpha},\hat{\lambda}),$$

respectively.

Thus, the logarithm of the maximum likelihood ratio is

$$T = \log\left(\frac{L_{WB}(\hat{\beta},\hat{\gamma})}{L_{PL}(\hat{\alpha},\hat{\lambda})}\right) = \log L_{WB}(\hat{\beta},\hat{\gamma}) - \log L_{PL}(\hat{\alpha},\hat{\lambda}).$$

If $T>0,$ the dataset is better fitted with the Weibull distribution, otherwise the PL distribution is selected.

4.2. Fisher information procedure

The Fisher information is beneficial to compute the asymptotic variances of various functions of estimators (Gupta and Kundu 2006). To obtain the asymptotic variances of the p^{th} percentile estimators, it required the function of p^{th} percentile estimators and the inverse of the Fisher information matrix.

The p^{th} $(0 < p < 1)$ percentile points of the $WB(\beta,\gamma)$ and $PL(\alpha,\lambda)$ distributions are

$$P_{WB} = \frac{1}{\gamma}(-\log(1-p))^{1/\beta}$$

and

$$P_{PL} = \left[-1-\frac{1}{\lambda}-\frac{1}{\lambda}W_{-1}\left(-(\lambda+1)(1-p)e^{-(\lambda+1)}\right)\right]^{1/\alpha}$$

respectively, where W_{-1} is the principal branch of the Lambert W function. The asymptotic variances of the p^{th} percentile estimators (Lawless 1982) of the Weibull and PL distributions are

$$V_{WB}(p) = \begin{bmatrix} \frac{\partial P_{WB}}{\partial \beta} & \frac{\partial P_{WB}}{\partial \gamma} \end{bmatrix} \begin{bmatrix} I_{WB11} & I_{WB12} \\ I_{WB21} & I_{WB22} \end{bmatrix}^{-1} \begin{bmatrix} \frac{\partial P_{WB}}{\partial \beta} \\ \frac{\partial P_{WB}}{\partial \gamma} \end{bmatrix},$$

where

$$\frac{\partial P_{WB}}{\partial \beta} = -\frac{1}{\gamma\beta^2}\left(-\log(1-p)\right)^{1/\beta}\log\left(-\log(1-p)\right),$$

$$\frac{\partial P_{WB}}{\partial \gamma} = -\frac{1}{\gamma^2}\left(-\log(1-p)\right)^{1/\beta}$$

and

$$V_{PL}(p) = \begin{bmatrix} \frac{\partial P_{PL}}{\partial \alpha} & \frac{\partial P_{PL}}{\partial \lambda} \end{bmatrix} \begin{bmatrix} I_{PL11} & I_{PL12} \\ I_{PL21} & I_{PL22} \end{bmatrix}^{-1} \begin{bmatrix} \frac{\partial P_{PL}}{\partial \alpha} \\ \frac{\partial P_{PL}}{\partial \lambda} \end{bmatrix},$$

where

$$\frac{\partial P_{PL}}{\partial \alpha} = -\frac{1}{\alpha^2}\left(-1-\frac{1}{\lambda}-\frac{1}{\lambda}W_{-1}(z)\right)^{1/\alpha}\log\left(-1-\frac{1}{\lambda}-\frac{1}{\lambda}W_{-1}(z)\right),$$

$$\frac{\partial P_{PL}}{\partial \lambda} = \frac{1}{\alpha}\left(-1-\frac{1}{\lambda}-\frac{1}{\lambda}W_{-1}(z)\right)^{(1-\alpha)/\alpha}\left(\frac{1}{\lambda^2}\left(1+W_{-1}(z)\right)-(1+p)e^{-(\lambda+1)}\frac{W_{-1}(z)}{z(1+W_{-1}(z))}\right),$$

where $z = -(\lambda+1)(1-p)e^{-(\lambda+1)}$.

We will use the asymptotic variance of the median estimator and the average asymptotic variance of the percentile estimators to select the appropriate distribution, defined as $V_{WB}(0.5), V_{PL}(0.5)$ and $\int_0^1 V_{WB}(p)\mathrm{d}p, \int_0^1 V_{WB}(p)\mathrm{d}p$ respectively. Moreover, the decision rule is to choose the distribution which has the less variance (Alshunnara et al. 2010; Gupta and Kundu 2003, 2006; Kundu and Raqab 2007)

5. Simulation study

In this section, we study the behaviors of the median estimator and the average asymptotic variance of the percentile estimators, based on the Monte Carlo simulations. In addition, we use sample size $n = 20, 50, 100, 200,$ and 1,000 repetitions. The median estimator and the average asymptotic variance of the percentile estimators are determined on 4 cases including:

1) $X \sim WB(\beta,1)$, where $\beta = 0.6, 0.8, 1.2, 1.6, 1.8$
2) $X \sim WB(1,\gamma)$, where $\gamma = 0.6, 0.8, 1.2, 1.6, 1.8$
3) $X \sim PL(\alpha,1)$, where $\alpha = 0.6, 0.8, 1.2, 1.6, 1.8$
4) $X \sim PL(1,\lambda)$, where $\lambda = 0.6, 0.8, 1.2, 1.6, 1.8$

Table 1 The number in the first row represents the asymptotic variances of the median estimator, and the numbers in the bracket show the average asymptotic variance of the percentile estimators, when a random variable X is distributed as $WB(\beta,1)$

β	n			
	20	50	100	200
0.6	0.1983	0.0724	0.0353	0.0173
	(10.1824)	(3.0857)	(1.3602)	(0.6447)
0.8	0.1305	0.0519	0.0259	0.0129
	(1.2948)	(0.4886)	(0.2345)	(0.1157)
1.2	0.0740	0.0304	0.0156	0.0078
	(0.2135)	(0.0840)	(0.0430)	(0.0212)
1.4	0.0603	0.0245	0.0123	0.0062
	(0.1312)	(0.0517)	(0.0259)	(0.0130)
1.6	0.0477	0.0201	0.0101	0.0051
	(0.0868)	(0.0359)	(0.0179)	(0.0089)
1.8	0.0404	0.0165	0.0084	0.0042
	(0.0657)	(0.0262)	(0.0133)	(0.0067)

Table 2 The number in the first row represents the asymptotic variances of the median estimator, and the numbers in the bracket show the average asymptotic variance of the percentile estimators, when a random variable X is distributed as $WB(1,\gamma)$

γ	n			
	20	50	100	200
0.6	0.2721	0.1102	0.0552	0.0275
	(1.2377)	(0.4646)	(0.2341)	(0.1156)
0.8	0.1553	0.0619	0.0310	0.0155
	(0.6841)	(0.2659)	(0.1311)	(0.0657)
1.2	0.0686	0.0277	0.0138	0.0069
	(0.3033)	(0.1203)	(0.0585)	(0.0289)
1.4	0.0506	0.0203	0.0101	0.0051
	(0.2260)	(0.0867)	(0.0432)	(0.0213)
1.6	0.0382	0.0155	0.0078	0.0039
	(0.1746)	(0.0667)	(0.0331)	(0.0164)
1.8	0.0302	0.0121	0.0061	0.0031
	(0.1451)	(0.0516)	(0.0259)	(0.0130)

Table 3 The number in the first row represents the asymptotic variances of the median estimator, and the numbers in the bracket show the average asymptotic variance of the percentile estimators, when a random variable X is distributed as $PL(\alpha,1)$

α	n			
	20	50	100	200
0.6	0.7647	0.3016	0.1517	0.0749
	(17.9853)	(5.4407)	(2.6739)	(1.2734)
0.8	0.3711	0.1473	0.0735	0.0367
	(2.1436)	(0.7986)	(0.3906)	(0.1897)
1.2	0.1379	0.0575	0.0289	0.0145
	(0.2978)	(0.1206)	(0.0597)	(0.0296)
1.4	0.0964	0.0400	0.0204	0.0103
	(0.1685)	(0.0681)	(0.0345)	(0.0174)
1.6	0.0715	0.0304	0.0152	0.0077
	(0.1092)	(0.0459)	(0.0228)	(0.0115)
1.8	0.0567	0.0233	0.0119	0.0060
	(0.0805)	(0.0325)	(0.0165)	(0.0083)

Table 4 The number in the first row represents the asymptotic variances of the median estimator, and the numbers in the bracket show the average asymptotic variance of the percentile estimators, when a random variable X is distributed as $PL(1,\lambda)$

λ	n			
	20	50	100	200
0.6	4.2382	4.3622	2.2061	1.1015
	(12.3866)	(12.8871)	(5.9882)	(3.0549)
0.8	2.2607	2.3728	1.1876	0.5968
	(6.7064)	(6.8783)	(3.3156)	(1.6553)
1.2	0.9510	0.9556	0.4900	0.2439
	(2.9725)	(2.8670)	(1.4428)	(0.7187)
1.4	0.6455	0.6789	0.3395	0.1714
	(2.0378)	(2.1036)	(1.0339)	(0.5171)
1.6	0.4863	0.5126	0.2498	0.1257
	(1.6620)	(1.6514)	(0.7810)	(0.3908)
1.8	0.3799	0.3838	0.1906	0.0958
	(1.3235)	(1.2700)	(0.6069)	(0.3049)

According to Tables 1-2, for the Weibull distribution, we discover that the $V_{WB}(0.5)$ and $\int_0^1 V_{WB}(p)\mathrm{d}p$ decrease when n, β, and γ increase.

When a random variable X has the PL distribution, there are decreasing trends of $V_{PL}(0.5)$ and $\int_0^1 V_{PL}(p)\mathrm{d}p$ if n, α, and λ increase, as shown in Tables 3-4.

In summary, the values of $V_{WB}(0.5)$, $V_{PL}(0.5)$, $\int_0^1 V_{WB}(p)\mathrm{d}p$, $\int_0^1 V_{PL}(p)\mathrm{d}p$ have decreasing trends when the shape and the scale parameters of the underlying distribution increase.

6. Applications

In this section, the Weibull and PL distributions are fitted to two real datasets. To select the appropriate distribution, we discuss primarily on the ratio of the maximized likelihoods, asymptotic variances of the median and average asymptotic variance of the percentile estimators. Other criteria including the Akaike Information Criterion (AIC), the Bayesian Information Criterion (BIC), and the Kolmogorov-Smirnov (KS) test are also presented.

Dataset 1 is the tensile strength of 69 carbon fibers tested under tension at gauge lengths of 20 mm. (Bader and Priest 1982) 1.312, 1.314, 1.479, 1.552, 1.700, 1.803, 1.861, 1.865, 1.944, 1.958, 1.966, 1.997, 2.006, 2.021, 2.027, 2.055, 2.063, 2.098, 2.14, 2.179, 2.224, 2.240, 2.253, 2.270, 2.272, 2.274, 2.301, 2.301, 2.359, 2.382, 2.382, 2.426, 2.434, 2.435, 2.478, 2.490, 2.511, 2.514, 2.535, 2.554, 2.566, 2.57, 2.586, 2.629, 2.633, 2.642, 2.648, 2.684, 2.697, 2.726, 2.770, 2.773, 2.800, 2.809, 2.818, 2.821, 2.848, 2.88, 2.954, 3.012, 3.067, 3.084, 3.090, 3.096, 3.128, 3.233, 3.433, 3.585, 3.585.

In addition, the Fisher information matrices are

$$\boldsymbol{I}_{WB}(\hat{\beta},\hat{\gamma}) = \begin{bmatrix} 0.0645 & 1.1937 \\ 1.1937 & 212.9763 \end{bmatrix},$$

$$\boldsymbol{I}_{PL}(\hat{\alpha},\hat{\lambda}) = \begin{bmatrix} 2.1041 & 39.7904 \\ 39.7904 & 809.1376 \end{bmatrix}.$$

The fitted results are shown in Table 5.

Table 5 Fitted results of dataset 1

	WB	PL
MLEs	$\hat{\beta} = 5.5055$	$\hat{\alpha} = 3.8670$
	$\hat{\gamma} = 0.3772$	$\hat{\lambda} = 0.0497$
$\log L$	-49.5961	-49.0595
AIC	103.1923	102.1190
BIC	107.6605	106.5873
KS test	0.0562	0.0443
p-value	0.9813	0.9990
$V(0.5)$	0.2802	0.2512
$\int_0^1 V(p)\mathrm{d}p$	0.3671	0.3622

According to Table 5, $T = -49.5961 + 49.0595 = -0.5366$, $V_{PL}(0.5) < V_{WB}(0.5)$, and $\int_0^1 V_{PL}(p)\mathrm{d}p < \int_0^1 V_{WB}(p)\mathrm{d}p$, which suggest choosing the PL distribution.

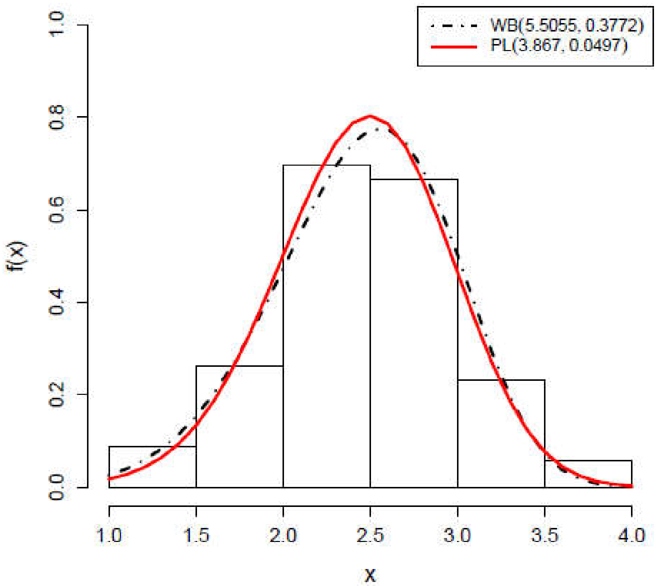

Figure 2 The histogram of the dataset 1 and the fitted density functions

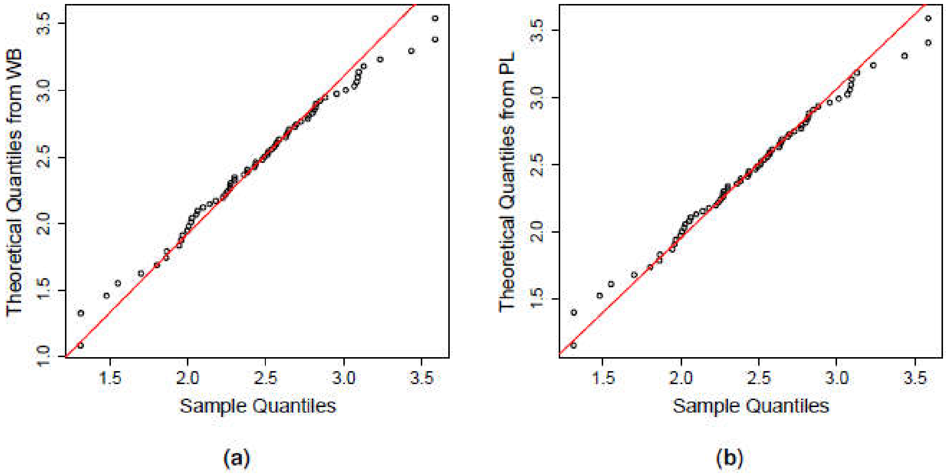

Figure 3 (a) Q-Q plot of fitted WB (b) Q-Q plot of fitted PL

Furthermore, the AIC and BIC of the PL distribution are less than those of the Weibull distribution, and the PL distribution has smaller KS statistic with larger p-value. Therefore, all of them also lead to the same conclusion.

Figures 2-3 show the fitted density functions and the quantile-quantile (Q-Q) plots, respectively, which illustrate that the fitted results based on the Weibull and PL distributions are quite similar to each other.

Dataset 2 is failure times of the air conditioning system of an airplane (Linhart and Zucchini 1986) 23, 261, 87, 7, 120, 14, 62, 47, 225, 71, 246, 21, 42, 20, 5, 12, 120, 11, 3, 14, 71, 11, 14, 11, 16, 90, 1, 16, 52, 95.

The fitted results are shown in Table 6.

Table 6 Fitted results of dataset 2

	WB	PL
MLEs	$\hat{\beta}=0.8535$	$\hat{\alpha}=0.6309$
	$\hat{\gamma}=0.0183$	$\hat{\lambda}=0.1635$
$\log L$	-151.9369	-151.9341
AIC	307.8738	307.8683
BIC	310.6762	310.6707
KS test	0.1532	0.1516
p-value	0.4817	0.4959
$V(0.5)$	2396.132	2140.871
$\int_0^1 V(p)\mathrm{d}p$	16192.3	18910.5

The associated Fisher information matrices are

$$\boldsymbol{I}_{WB}(\hat{\beta},\hat{\gamma})=\begin{bmatrix}2.6194 & 24.6257\\ 24.6256 & 2185.1381\end{bmatrix},$$

$$\boldsymbol{I}_{PL}(\hat{\alpha},\hat{\lambda})=\begin{bmatrix}38.0470 & 48.8067\\ 48.8067 & 74.0767\end{bmatrix}.$$

With respect to Table 6, $T=-151.9369+151.9341=-\ 0.0028$ and $V_{PL}(0.5)<V_{WB}(0.5)$ imply that the PL distribution fits the given data better than the Weibull distribution. Additionally, the same conclusion is made by considering the AIC, BIC, and the KS test. However, $\int_0^1 V_{PL}(p)\mathrm{d}p>\int_0^1 V_{WB}(p)\mathrm{d}p,$ which indicates that the variation of the pth percentile estimators of the PL distribution is greater than that of the Weibull distribution. In other words, considering overall percentiles, the Weibull distribution can fit dataset 2 better than the PL distribution, especially for percentiles on the tail as shown by Q-Q plot in Figure 5. In further work, one can include other distributions to be evaluated performance to fit this dataset.

Finally, Figures 4-5 show the fitted density functions and Q-Q plots respectively.

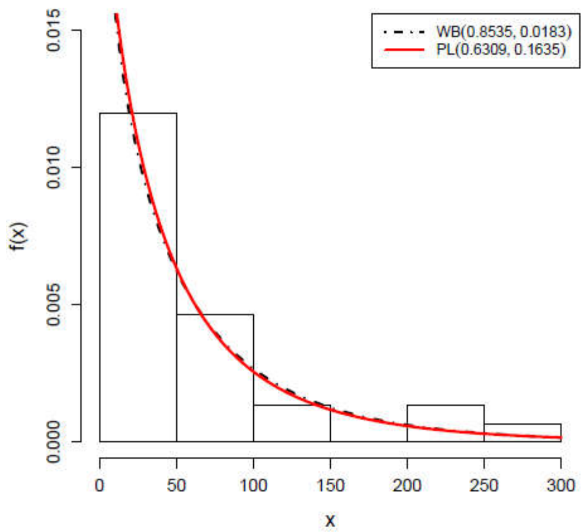

Figure 4 The histogram of the dataset 2 and the fitted density functions

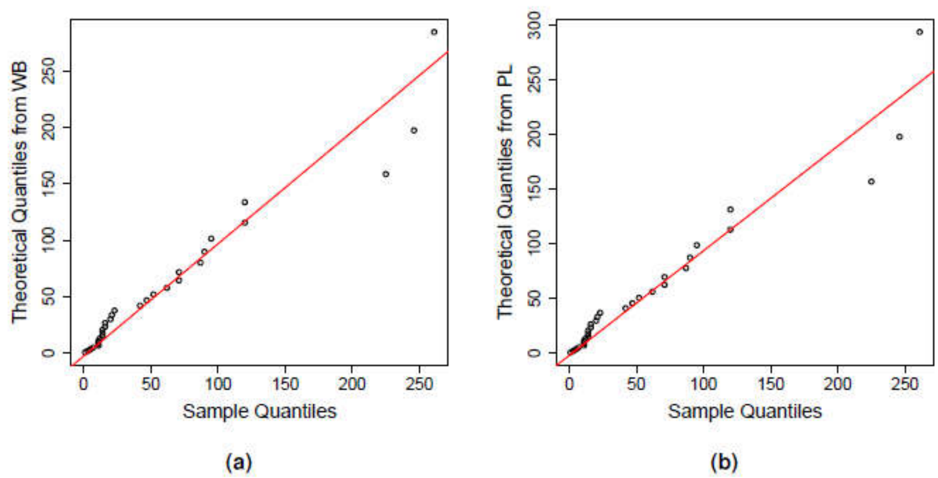

Figure 5 (a) Q-Q plot of fitted WB (b) Q-Q plot of fitted PL

7. Conclusions

We study the procedures for selecting the most appropriate distribution based on the likelihood function and the Fisher information matrix. The two procedures are applied to discriminate between the Weibull and PL distributions. In addition, the Fisher information matrix is used to compute the asymptotic variance of the pth percentile estimators including the asymptotic variances of the median and average asymptotic variance of the percentile estimators. Indeed, the asymptotic variance of the p^{th} percentile estimators is useful as it provides the further information of the Q-Q plot, especially the variation. In application, we consider the values of the likelihood function and the asymptotic variance of the pth percentile estimators as the criteria of discrimination. We discover the different and indifferent conclusions they lead to. In fact, in some cases, it is quite difficult to say that the dataset belongs to the Weibull distribution or the PL distribution. Moreover, ones can get the further information to choose the best model by applying the Fisher information matrix to compute the variance of other estimators.

Acknowledgements

The authors are grateful to Department of Statistics, Faculty of Science, Kasetsart University and Graduate School Kasetsart University.

References

Alshunnara FS, Raqaba MZ, Kundu D. On the comparison of the Fisher Information of the log-normal and generalized Rayleigh distributions. J. Appl. Stat. 2010; 37: 391-404.

Bader MG, Priest AM. Statistical aspects of fiber and bundle strength in hybrid composites. In: Hayashi T, Kawata K, Umekawa S, editors. Progress in Science and Engineering Composites. 1982; 1: 1129-1136.

Efron B, Johnstone I. Fisher information in terms of the hazard rate. Ann. Stat. 1990; 18: 38-62.

Ghitany ME, Al-Mutairi DK, Balakrishnan N, Al-Enezi LJ. Power Lindley distribution and associated inference. Comput. Stat. Data Anal. 2013; 64: 20-33.

Gupta RD, Kundu D. Discriminating between Weibull and generalized exponential distributions. Comput. Stat. Data Anal. 2003; 43: 179-196.

Gupta RD, Kundu D. On the comparison of Fisher information of the Weibull and GE distributions. J. Stat. Plann. Infer. 2006; 136: 3130-3144.

Hofmann G. Comparing the Fisher information in record data and random observations. Stat. Paper. 2004; 45: 517-528.

Hofmann G, Balakrishnan N, Ahmadi J. Characterization of hazard function factorization by Fisher information in minima and upper record values. Stat. Probab. Lett. 2005; 72: 51-57.

Kundu D, Raqab MZ. Discriminating between the generalized Rayleigh and log-normal distribution. Statistics. 2007; 41: 505-515.

Lawless JF. Statistical Models and Methods for Lifetime Data. New York. Wiley. 1982.

Lehmann EL. Theory of Point Estimation. New York. Wiley. 1991.

Linhart H, Zucchini W. Model Selection. New York. Wiley. 1986.

Nadarajah S. Information matrices for Laplace and Pareto mixtures. Comput Stat Data Anal. 2006; 50: 950-966.

Pakyari R. Discriminating between generalized exponential, geometric extreme exponential and Weibull distributions. J. Stat. Comput. Simulat. 2010; 80: 1403-1412.

R Core Team. R: A Language and Environment for Statistical Computing. Vienna. Austria. 2015. Available from URL: http://www.R-project.org/.

Raqaba MZ. Discriminating between the generalized Rayleigh and Weibull distributions. J. Appl. Stat. 2013; 40: 1480-1493.

Rinne H. The Weibull Distribution: A Handbook. Chapman and Hall/CRC. 1986.