The Poisson Nadarajah-Haghighi Distribution: Different Methods of Estimation

Sajid Ali[1,*], Sanku Dey[2], M. H. Tahir[3], and Muhammad Mansoor[4]

[1]*Department of Statistics, Quaid-i-Azam University, Islamabad 45320, Pakistan*
[2]*Department of Statistics, St. Anthony's College, Shillong 793001, India*
[3]*Department of Statistics, The Islamia University of Bahawalpur, Bahawalpur 63100, Pakistan*
[4]*Department of Statistics, Government Sadiq Egerton College, Bahawalpur, Pakistan*
E-mail: sajidali.qau@hotmail.com
[*]*Corresponding Author*

Received 15 March 2021; Accepted 05 July 2021;
Publication 23 August 2021

Abstract

Estimation of parameters of Poisson Nadarajah-Haghighi (PNH) distribution from the frequentist and Bayesian point of view is discussed in this article. To this end, we briefly described ten different frequentist approaches, namely, the maximum likelihood estimators, percentile based estimators, least squares estimators, weighted least squares estimators, maximum product of spacings estimators, minimum spacing absolute distance estimators, minimum spacing absolute-log distance estimators, Cramér-von Mises estimators, Anderson-Darling estimators and right-tail Anderson-Darling estimators. To assess the performance of different estimators, Monte Carlo simulations are done for small and large samples. The performance of the estimators is compared in terms of their bias, root mean squares error, average absolute difference between the true and estimated distribution functions, and the maximum absolute difference between the true and estimated distribution functions of the estimates using simulated data. For the Bayesian inference of the

Journal of Reliability and Statistical Studies, Vol. 14, Issue 2 (2021), 415–450.
doi: 10.13052/jrss0974-8024.1423

unknown parameters, we use Metropolis–Hastings (MH) algorithm to calculate the Bayes estimates and the corresponding credible intervals. Results from the simulation study suggests that among the considered classical methods of estimation, weighted least squares and the maximum product spacing estimators uniformly produces the least biases of the estimates with least root mean square errors. However, Bayes estimates perform better than all other estimates. Finally, we discuss a practical data set to show the application of the distribution.

Keywords: Exponential distribution, hazard rate, lifetime data, maximum likelihood method, Bayesian estimation, Nadarajah-Haghighi distribution, Poisson distribution.

1 Introduction

Although there are many continuous and discrete distributions in statistics literature, the exponential distribution enjoys a special place due to its memory-less and constant hazard rate properties. Thus, it is used as a benchmark model in the reliability analysis. To overcome constant hazard rate, many extensions of the exponential distribution have been introduced in the literature, for example, exponentiated-exponential (EE) (Gupta and Kundu, 1999) and beta-exponential (BE) (Nadarajaha and Kotz, 2006), among many others. Nadarajah and Haghighi (2011) introduced a new extension of the exponential and to define it, let Z have the Nadarajah-Haghighi (NH for short) distribution, say $Z \sim \text{NH}(\alpha, \lambda)$. The cumulative distribution function (cdf) of NH distribution is given by

$$G(x) = 1 - \mathrm{e}^{1-(1+\lambda\, x)^{\alpha}}, \tag{1}$$

where $\lambda > 0$ is the scale parameter and $\alpha > 0$ is the shape parameter. The NH distribution reduces to exponential distribution assuming $\alpha = 1$. The probability density function (pdf) corresponding to (1) is given by

$$g(x) = \alpha\lambda\; (1 + \lambda\, x)^{\alpha-1}\, \mathrm{e}^{1-(1+\lambda x)^{\alpha}}, \quad x > 0. \tag{2}$$

Nadarajah and Haghighi (2011) pointed out that the density function (2) always has zero mode. Additionally, the hazard rate function (hrf) of the NH distribution can be increasing, decreasing, and constant. It is noted by Nadarajah and Haghighi (2011) that the NH density function can be monotonically decreasing and yet increasing hrf. Also, if Y is a Weibull random variable with the shape parameter α and scale parameter λ, then the density (2) has the same as that of the random variable $Z = Y - \lambda^{-1}$ truncated at zero, i.e., the NH distribution can be interpreted as the truncated Weibull distribution.

Recently, Mansoor et al. (2020a) proposed the Poisson Nadarajah-Haghighi (PNH) model to model reliability systems. To this end, consider a company formed by N systems functioning independently at a given time, where N is a zero-truncated Poisson (ZTP) random variable (rv) with the probability mass function (pmf)

$$\mathbb{P}(N=n) = \frac{\theta^n}{n!\,(\mathrm{e}^{\theta}-1)}, \quad n = 1, 2, \ldots .$$

Next, suppose that each system consists of β parallel units. The system will fail if all units fail and assume that the failure times of the units for the ith system, say $Z_{i,1}, \ldots, Z_{i,\beta}$ are independent and identically NH random variables with scale parameter λ and shape parameter α. Let Y_i denote the failure time of the ith system and X represents the time to failure of the first of the N functioning systems. Then, one can write $X = \min(Y_1, \ldots, Y_N)$ and the conditional cdf of X given N is

$$\begin{aligned} F(X \mid N) &= 1 - [\mathbb{P}\,(Y_1 > t)]^N = 1 - [1 - \mathbb{P}\,(Z_{1,1} \le t, \ldots, Z_{1,\beta} \le X)]^N \\ &= 1 - \left[1 - \left\{1 - \mathrm{e}^{1-(1+\lambda\,x)^{\alpha}}\right\}^{\beta}\right]^N . \end{aligned}$$

Hence, the unconditional cdf of X is given by

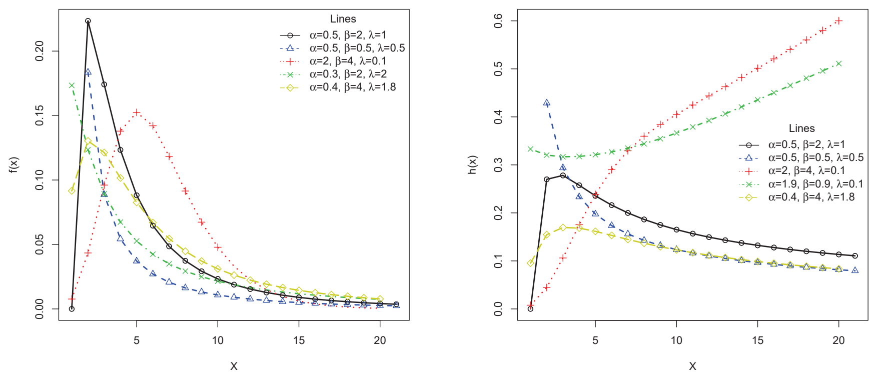
$$F(x) = \frac{1 - \exp\left\{ -\theta \left[1 - \mathrm{e}^{1-(1+\lambda\,x)^{\alpha}}\right]^{\beta} \right\}}{1 - \mathrm{e}^{-\theta}}.$$

For simplicity, let $\theta = 1$. Then,

$$F(x) = \frac{1 - \exp\left(-\{1 - \mathrm{e}^{1-(1+\lambda\,x)^{\alpha}}\}^{\beta}\right)}{1 - \mathrm{e}^{-1}}. \tag{3}$$

and the pdf corresponding to (3) is given by

$$\begin{aligned} f(x) = {} & \frac{\beta\alpha\lambda\,(1+\lambda\,x)^{\alpha-1}\,\mathrm{e}^{1-(1+\lambda\,x)^{\alpha}}\{1 - \mathrm{e}^{1-(1+\lambda\,x)^{\alpha}}\}^{\beta-1}}{1 - \mathrm{e}^{-1}} & (4) \\ & \times \exp\left(-\{1 - \mathrm{e}^{1-(1+\lambda\,x)^{\alpha}}\}^{\beta}\right). & (5) \end{aligned}$$



(a) (b)

Figure 1 Plots of the PNH (a) density and (b) hazard rate for some selected parameter values.

Hereafter, a random variable X with cdf (3) is called the *Poisson Nadarajah-Haghighi* (PNH) distribution and denoted by $X \sim \mathrm{PNH}(\beta, \alpha, \lambda)$. Clearly, if $\alpha = 1$, the PNH distribution reduces to the Poisson-exponential (PE) distribution. This distribution is introduced by Mansoor et al. (2020a), however, many properties, especially a comparison of different estimation methods has not been considered in the literature. The survival function (sf) and hrf of X are given by

$$S(x) = \frac{\exp\left(-\{1 - \mathrm{e}^{1-(1+\lambda\, x)^{\alpha}}\}^{\beta}\right) - \mathrm{e}^{-1}}{1 - \mathrm{e}^{-1}} \tag{6}$$

and

$$h(x) = \frac{\beta\, \alpha\, \lambda\, (1 + \lambda\, x)^{\alpha-1}\, \mathrm{e}^{1-(1+\lambda\, x)^{\alpha}}\, \{1 - \mathrm{e}^{1-(1+\lambda\, x)^{\alpha}}\}^{\beta-1}}{1 - \exp\left(-\left[1 - \{1 - \mathrm{e}^{1-(1+\lambda\, x)^{\alpha}}\}^{\beta}\right]\right)} \tag{7}$$

respectively. Figures 1(a) and 1(b) display some plots of the density and hrf of X for different values of α, β and λ. Figure 1(a) reveals that the PNH density has decreasing and unimodal (right-skewed) shape, whereas Figure 1(b) indicates that the PNH hazard rate is decreasing, increasing, bathtub (BT) and up-side bathtub (UBT).

Parameter estimation plays a vital role in statistics and the maximum likelihood estimation is generally a starting point to estimate parameters. The popularity of this method is due to its simple and intuitive formulation. For example, estimators obtained by this method are asymptotically consistent and normally distributed (Lehmann and Casella, 2003). However, there are other estimation methods in the literature, which are commonly used. For example, Kundu and Raqab (2005) for generalized Rayleigh distributions, Teimouri et al. (2013) for Weibull distribution, Ali et al. (2020b) for two-parameter logistic-exponential distribution, Dey et al. (2014, 2015, 2016, 2017b,a,c,d) for the two-parameter Rayleigh, weighted exponential, two-parameter Maxwell, exponentiated-Chen, Dagum, transmuted-Rayleigh, two parameter exponentiated-Gumbel, new extension of generalized exponential and NH distributions. Recent literature in this direction may be seen in Alizadeh et al. (2020), Eliwa et al. (2020), Tahir et al. (2018), Ali et al. (2020c), Mansoor et al. (2020b), Ali et al. (2020a), Shafqat et al. (2020) and references cited therein. These methods are the method of moment estimation, method of L-moment estimation, method of probability weighted moment estimation, method of least-squares estimation, method of weighted least-square estimation, method of maximum product spacing estimation and method of minimum distance estimation and so on.

The aim of this study is to provide a comprehensive comparison of different frequentist methods of estimation for the PNH distribution. To this end, we assume different sample sizes and different combination of parameter values. We focus on the maximum likelihood estimators, percentile based estimators, maximum product of spacings estimators, least-squares estimators, weighted least-squares estimators, Cramér-von-Mises estimators, Anderson-Darling estimators and right-tail Anderson-Darling estimators. As it is difficult to compare the performances of different methods theoretically, extensive simulations are performed to compare the performances of the different estimators based on their relative bias, root mean squares error, the average absolute difference between the true and estimated distribution functions, and the maximum absolute difference between the true and estimated distribution functions of the estimates. The originality of this study comes from the fact that there has been no previous work comparing all of these estimation methods for the PNH distribution. Further, we also consider the Bayesian estimation of the unknown parameters under the assumptions of independent gamma priors on the scale and shape parameters, respectively. We present a Metropolis-Hastings (MH) algorithm to compute the Bayes estimates and the associated credible intervals. A real life data set is also analyzed for illustrative purposes. Thus, the study will be a guideline for choosing the best estimation method for the PNH distribution, which we think would be interesting for applied statisticians.

The rest of the paper is organized as follows. Section 2 presents the quantile function, moments and shapes of the pdf and hrf for the proposed model. Section 3 describes ten different frequentist methods of estimation. In Section 4, a simulation study is carried out to compare the performance of different methods of estimation for the proposed model. In Section 5, Bayesian analysis is conducted using the Metropolis-Hastings (MH) algorithm. In Section 6, the usefulness of the PNH distribution is illustrated using a real dataset. Finally, Section 7 offers some concluding remarks.

2 Basic Statistical Properties of PNH Distribution

This section discusses some basic statistical properties of the PNH distribution.

2.1 Quantile function

To generate random variables from the PNH distribution, we invert Equation (3) as $X = F^{-1}(u)$, where $u \sim \text{Uniform}(0,1)$. The explicit form of the PNH quantile is

$$\begin{aligned} X &= F^{-1}(u) \\ &= \frac{1}{\lambda}\left[\left\{1 - \ln\left(1 - \left(-\ln\left\{1 - u\left(1 - \exp(-1)\right)\right\}\right)^{1/\beta}\right)\right\}^{1/\alpha} - 1\right]. \end{aligned} \quad (8)$$

Further, the quantile function can be used to investigate the skewness and kurtosis measures. For example, the Bowley skewness (Kenney and Keeping, 1962) based on quantiles is given by

$$B = \frac{F^{-1}(3/4) + F^{-1}(1/4) - 2\,F^{-1}(2/4)}{F^{-1}(3/4) - F^{-1}(1/4)}.$$

Similarly, the Moors' kurtosis (Moors, 1988) is

$$M = \frac{F^{-1}(3/8) - F^{-1}(1/8) + F^{-1}(7/8) - F^{-1}(5/8)}{F^{-1}(6/8) - F^{-1}(2/8)}.$$

2.2 Moments

Many properties of a distribution can be studied using moments, e.g., tendency, dispersion, skewness, and kurtosis. The nth moment expression of

PNH is given by

$$\mu'_n = \frac{\beta}{\lambda^n(1-\exp(-1))} \sum_{i,j=0}^{\infty} \sum_{k=0}^{n} \frac{(-1)^{i+n-k}\exp(i+1)}{(i+1)^{k/2+1}} \binom{\beta(i+1)-1}{j}\binom{n}{k}\Gamma\left(k/2+1, i+1\right),$$

where $\Gamma(a, x)$ denotes the incomplete gamma function defined as $\Gamma(a, x) = \int_x^{\infty} t^{a-1}\exp(-t)dt$.

The graphical depiction of the mean, variance, skewness, and kurtosis is given in Figure 2. It is noticed that the mean and variance decreased by increasing λ while increased by increasing β. Similarly, the skewness and kurtosis decreased by decreasing β, and λ is not significant as observed in the cases of mean and variance.

2.3 Shapes of the Density and Hazard Rate Functions of PNH Model

To study the shapes of the density and the hrf, we determine critical points of the PNH density by $\partial \ln f(x)/\partial x = 0$, which are the roots of the following equation.

$$\frac{(\alpha-1)\lambda}{(1-\ln w)^{1/\alpha}} - \alpha\lambda(1-\ln w)^{1-1/\alpha} + \frac{(\beta-1)\lambda\alpha(\ln w)^{1-1/\alpha}w}{1-w} - \alpha\beta\lambda(1-\ln w)^{1-1/\alpha}w(1-w)^{\beta-1} = 0. \tag{9}$$

The critical points of the PNH hrf can be obtained from the following equation

$$\frac{(\alpha-1)\lambda}{(1-\ln w)^{1/\alpha}} - \alpha\lambda(1-\ln w)^{1-1/\alpha} + \frac{(\beta-1)\lambda\alpha(\ln w)^{1-1/\alpha}w}{1-w} - \frac{\alpha\beta\lambda(1-\ln w)^{1-1/\alpha}w(1-w)^{\beta-1}\exp\big(-1-(1-w)^{\beta}\big)}{1-\exp\big(-(1-(1-w)^{\beta})\big)} = 0, \tag{10}$$

where $w = \exp\big(1-(1+\lambda x)^{\alpha}\big)$. One can examine numerically the local maximum, minimum and inflexion points of Equations (9) and (10).

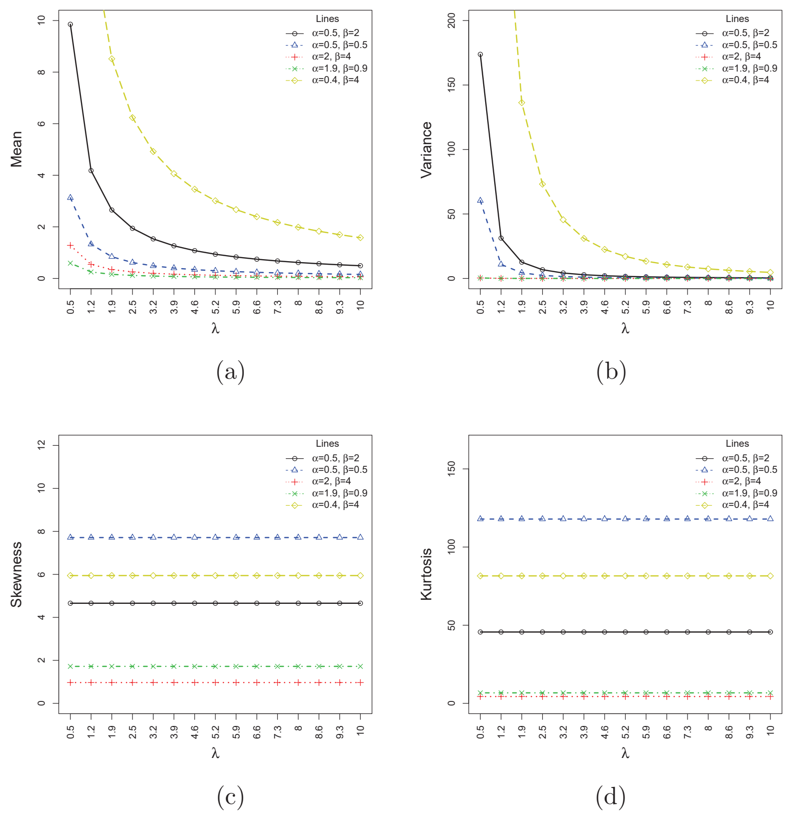


Figure 2 Plots of the PNH (a) Mean (b) Variance (c) Skewness, and (d) Kurtosis for some selected parameter values.

Another property to characterize the distribution is the log-concave, i.e., the density is log-concave if $d^2/dx^2 \log f < 0$, otherwise convex. The hazard would be decreasing if the density is log-concave. For the PNH, it is observed that the density is log-concave for $\alpha > 1$ with a fixed λ. Moreover, for $\beta > \alpha$ the density is also observed log-concave. Similarly, hazard rate average $HRA(x) = \frac{H(x)}{x} = \frac{1}{x}\int_0^x h(u)du$ can be used to characterize the distribution whether it is increasing (decreasing) hazard rate IDHR (DIHR) if $\frac{d}{dx}\frac{H(x)}{x} \geq 0 (\leq 0)$ for $x > 0$. The PNH is DIHR for $\beta > 1, \forall \quad \alpha, \lambda$.

A density is said to be new-better-than-used (NBU) if $\Delta(x, y) = \frac{S(x)S(y)}{S(x+y)} \geq 1$, for $x, y \geq 0$, otherwise new-worse-than-used (NWU). From Figure 3, it is clear that the PNH is NBU for $\alpha > 1$.

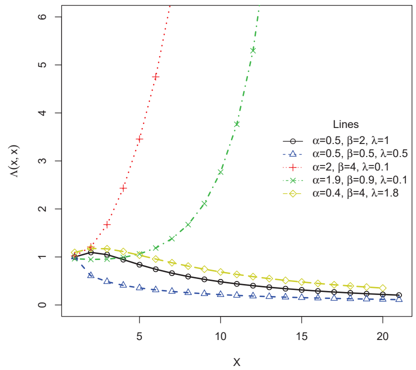


Figure 3 Plot to identify NBU and NWU of the PNH for some selected parameter values.

3 Methods of Estimation

This section describes ten classical methods for estimating the parameters, α, β and λ assuming $x = (x_1, \ldots, x_n)$ a random sample of size n from the distribution (4) with unknown parameters α, β and λ.

3.1 Method of Maximum Likelihood

It is well-known that the method of maximum likelihood is the most popular method in statistical inference, since it has several attractive properties (Lehmann and Casella, 2003). $\boldsymbol{\theta} = (\beta, \alpha, \lambda)^\top$. The log-likelihood for $\boldsymbol{\theta} = (\beta, \alpha, \lambda)^\top$ based on a given sample is given by

$$\begin{aligned}\ell(\boldsymbol{\theta}) &= n\log(\beta\,\alpha\,\lambda) - n\log(1-\mathrm{e}^{-1}) + (\alpha-1)\sum_{i=1}^{n}\log(1+\lambda x_i)\\ &\quad + \sum_{i=1}^{n}(1-(1+\lambda x_i)^{\alpha})\\ &\quad + (\beta-1)\sum_{i=1}^{n}\log\left[1-\mathrm{e}^{1-(1+\lambda x_i)^{\alpha}}\right] - \sum_{i=1}^{n}\left[1-\mathrm{e}^{1-(1+\lambda x_i)^{\alpha}}\right]^{\beta}.\end{aligned} \tag{11}$$

The maximum likelihood estimators (MLEs) of the model parameters can be obtained by maximizing the log-likelihood function $\ell(\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$. There are several routines available for numerical maximization of (11) given in the `R` program (`optim` function), `SAS` (`PROC NLMIXED`), `Ox` (sub-routine

MaxBFGS). Alternatively, one can differentiate (11) and solve the resulting nonlinear likelihood equations.

The partial derivatives of $\ell(\boldsymbol{\theta})$ with respect to the parameters are given by

$$\frac{\partial\,\ell(\boldsymbol{\theta})}{\partial\,\beta} = \frac{n}{\beta} + \sum_{i=1}^{n} \log\left[1 - \mathrm{e}^{1-(1+\lambda x_i)^{\alpha}}\right] - \sum_{i=1}^{n}\left[1 - \mathrm{e}^{1-(1+\lambda x_i)^{\alpha}}\right]^{\beta} \log\left[1 - \mathrm{e}^{1-(1+\lambda x_i)^{\alpha}}\right],$$

$$\frac{\partial\,\ell(\boldsymbol{\theta})}{\partial\,\alpha} = \frac{n}{\alpha} + \sum_{i=1}^{n} \log[1-(1+\lambda x_i)] + (\beta-1)\sum_{i=1}^{n}\frac{\mathrm{e}^{1-(1+\lambda x_i)^{\alpha}}}{1-\mathrm{e}^{1-(1+\lambda x_i)^{\alpha}}}\left[(1+\lambda x_i)^{\alpha}\,\log(1+\lambda x_i)\right] - \beta\sum_{i=1}^{n}(1+\lambda x_i)^{\alpha}\,\log(1+\lambda x_i)\,\mathrm{e}^{1-(1+\lambda x_i)^{\alpha}} \left[1 - \mathrm{e}^{1-(1+\lambda x_i)^{\alpha}}\right]^{\beta-1} - \sum_{i=1}^{n}(1+\lambda x_i)^{\alpha}\log(1+\lambda x_i),$$

$$\frac{\partial\,\ell(\boldsymbol{\theta})}{\partial\,\lambda} = \frac{n}{\lambda} + (\alpha-1)\sum_{i=1}^{n}\frac{x_i}{1+\lambda x_i} + \alpha(\beta-1)\sum_{i=1}^{n} \frac{x_i(1+\lambda x_i)^{\alpha-1}\mathrm{e}^{1-(1+\lambda x_i)^{\alpha}}}{1-\mathrm{e}^{1-(1+\lambda x_i)^{\alpha}}} - \alpha\,\beta\sum_{i=1}^{n} x_i(1+\lambda x_i)^{\alpha-1}\mathrm{e}^{1-(1+\lambda x_i)^{\alpha}}\left[1 - \mathrm{e}^{1-(1+\lambda x_i)^{\alpha}}\right] - \alpha\sum_{i=1}^{n}(1+\lambda x_i)^{\alpha-1}.$$

The MLE $\hat{\boldsymbol{\theta}} = (\hat{\beta}, \hat{\alpha}, \hat{\lambda})^{\top}$ of $\boldsymbol{\theta} = (\beta, \alpha, \lambda)^{\top}$ can be obtained by solving simultaneously the following normal equations

$$\frac{\partial\,\ell(\boldsymbol{\theta})}{\partial\,\beta} = 0, \frac{\partial\,\ell(\boldsymbol{\theta})}{\partial\,\alpha} = 0, \frac{\partial\,\ell(\boldsymbol{\theta})}{\partial\,\lambda} = 0.$$

There is no closed-form expressions for $\hat{\beta}, \hat{\alpha}$ and $\hat{\lambda}$ and therefore numerical computations using nonlinear optimization algorithm, such as the Newton-Raphson iterative method, should be used.

3.2 Method of Maximum and Minimum Spacing Distance Estimators

Cheng and Amin (1979) introduced the maximum product of spacings (MPS) method as an alternative to MLE for the estimation of parameters of continuous univariate distributions. Ranneby (1984) independently developed the same method as an approximation to the Kullback-Leibler measure of information. Let $F(x_{(j)})$ denote the distribution function of the ordered random variables $x_{(1)} < x_{(2)} < \cdots < x_{(n)}$, where $\{x_1, x_2, \cdots, x_n\}$ is a random sample of size n from the cdf $F(\cdot)$.

Let define the uniform spacings of a random sample from the PNH distribution distribution as

$$D_i(\alpha, \beta, \lambda) = F\left(x_{i:n} \mid \alpha, \beta, \lambda\right) - F\left(x_{i-1:n} \mid \alpha, \beta, \lambda\right),$$

where $F(x_{0:n} \mid \alpha, \beta, \lambda) = 0$ and $F(x_{n+1:n} \mid \alpha, \beta, \lambda) = 1$. Clearly, $\sum_{i=1}^{n+1} D_i(\alpha, \beta, \lambda) = 1$.

Following Cheng and Amin (1983), the maximum product of spacings estimators $\widehat{\alpha}_{MPS}, \widehat{\beta}_{MPS}$ and $\widehat{\lambda}_{MPS}$ of the parameters α, β and λ are obtained by maximizing the geometric mean of the spacings with respect to α, β and λ

$$G(\alpha, \beta, \lambda) = \left[\prod_{i=1}^{n+1} D_i(\alpha, \beta, \lambda)\right]^{\frac{1}{n+1}}, \tag{12}$$

or, equivalently, by maximizing the function

$$H(\alpha, \beta, \lambda) = \frac{1}{n+1} \sum_{i=1}^{n+1} \log D_i(\alpha, \beta, \lambda). \tag{13}$$

The estimators $\widehat{\alpha}_{MPS}$, $\widehat{\beta}_{MPS}$ and $\widehat{\lambda}_{MPS}$ of the parameters α, β and λ can also be obtained by solving the nonlinear equations

$$\frac{\partial}{\partial \alpha} H\left(\alpha, \beta, \lambda\right) = \frac{1}{n+1} \sum_{i=1}^{n+1} \frac{1}{D_i(\alpha, \beta, \lambda)} \left[\Delta_1(x_{i:n} | \alpha, \beta, \lambda) - \Delta_1(x_{i-1:n}) | \alpha, \beta, \lambda)\right] = 0, \tag{14}$$

$$\frac{\partial}{\partial\beta}H\left(\alpha,\beta,\lambda\right)=\frac{1}{n+1}\sum_{i=1}^{n+1}\frac{1}{D_i(\alpha,\beta,\lambda)}$$
$$[\Delta_2(x_{i:n}|\alpha,\beta,\lambda)-\Delta_2(x_{i-1:n})|\alpha,\beta,\lambda]=0, \qquad (15)$$
$$\frac{\partial}{\partial\lambda}H\left(\alpha,\beta,\lambda\right)=\frac{1}{n+1}\sum_{i=1}^{n+1}\frac{1}{D_i(\alpha,\beta,\lambda)}$$
$$[\Delta_3(x_{i:n})|\alpha,\beta,\lambda)-\Delta_3(x_{i-1:n})|\alpha,\beta,\lambda)]=0, \qquad (16)$$

where

$$\Delta_1\left(x_{i:n}\mid\alpha,\beta,\lambda\right)=\frac{1}{1-e^{-1}}[\beta e^{-[1-e^{1-(1+\lambda x_{i:n})^{\alpha}}]^{\beta}}(1+\lambda x_{i:n})^{\alpha}$$
$$\log(1+\lambda x_{i:n})e^{1-(1+\lambda x_{i:n})^{\alpha}}(1-e^{1-(1+\lambda x_{i:n})^{\alpha}})^{\beta-1}], \qquad (17)$$

$$\Delta_2\left(x_{i:n}\mid\alpha,\beta,\lambda\right)=\frac{1}{1-e^{-1}}[e^{-[1-e^{1-(1+\lambda x_{i:n})^{\alpha}}]^{\beta}}(1-e^{1-(1+\lambda x_{i:n})^{\alpha}})^{\beta}$$
$$\log(1-e^{1-(1+\lambda x_{i:n})^{\alpha}})] \qquad (18)$$

and

$$\Delta_3\left(x_{i:n}\mid\alpha,\beta,\lambda\right)=\frac{1}{1-e^{-1}}[\alpha\beta x_{i:n}(1+\lambda x_{i:n})^{\alpha-1}e^{-[1-e^{1-(1+\lambda x_{i:n})^{\alpha}}]^{\beta}}$$
$$(1-e^{1-(1+\lambda x_{i:n})^{\alpha}})^{\beta-1}e^{1-(1+\lambda x_{i:n})^{\alpha}})]. \qquad (19)$$

Cheng and Amin (1983) showed that maximizing H as a method of parameter estimation is as efficient as the MLE estimation. Further, the MPS estimators are also consistent under more general conditions than the MLE estimators.

Similarly, the minimum spacing distance estimators of $\widehat{\alpha}_{MSADE}$, $\widehat{\beta}_{MSADE}$ and $\widehat{\lambda}_{MSADE}$ of α, β and λ are obtained by minimizing

$$T\left(\alpha,\beta,\lambda\right)=\sum_{i=1}^{n+1}h\left(D_i\left(\alpha,\beta,\lambda\right),\frac{1}{n+1}\right), \qquad (20)$$

where $h(x,y)$ is an appropriate distance. Some choices of $h(x,y)$ are $|x-y|$ and $|\log x-\log y|$, which are called absolute and absolute-log distance, respectively. These estimators are called the minimum spacing absolute distance estimator (MSADE) and the minimum spacing absolute-log distance estimator (MSALDE). This method was originally proposed by Torabi (2008). The MSADE and MSALDE of parameters α, β and λ can be obtained by minimizing

$$T\left(\alpha, \beta, \lambda\right) = \sum_{i=1}^{n+1} \left|(D_i(\alpha, \beta, \lambda) - \frac{1}{n+1}\right| \tag{21}$$

and

$$T\left(\alpha, \beta, \lambda\right) = \sum_{i=1}^{n+1} \left|\log D_i(\alpha, \beta, \lambda) - \log \frac{1}{n+1}\right|, \tag{22}$$

with respect to α, β and λ respectively.

The estimators $\hat{\alpha}_{MSADE}$, $\hat{\beta}_{MSADE}$ and $\hat{\lambda}_{MSADE}$ of α, β and λ can be obtained by solving the following nonlinear equations

$$\frac{\partial}{\partial \alpha} T(\alpha, \beta, \lambda) = \sum_{i=1}^{n+1} \frac{D_i(\alpha, \beta, \lambda) - \frac{1}{n+1}}{\left|D_i(\alpha, \beta, \lambda) - \frac{1}{n+1}\right|}$$

$$\left[\Delta_1\left(x_{i:n} \mid \alpha, \beta, \lambda\right) - \Delta_1\left(x_{i-1:n} \mid \alpha, \beta, \lambda\right)\right] = 0$$

$$\frac{\partial}{\partial \beta} T(\alpha, \beta, \lambda) = \sum_{i=1}^{n+1} \frac{D_i(\alpha, \beta, \lambda) - \frac{1}{n+1}}{\left|D_i(\alpha, \beta, \lambda) - \frac{1}{n+1}\right|}$$

$$\left[\Delta_2\left(x_{i:n} \mid \alpha, \beta, \lambda\right) - \Delta_2\left(x_{i-1:n} \mid \alpha, \beta, \lambda\right)\right] = 0,$$

$$\frac{\partial}{\partial \lambda} T(\alpha, \beta, \lambda) = \sum_{i=1}^{n+1} \frac{D_i(\alpha, \beta, \lambda) - \frac{1}{n+1}}{\left|D_i(\alpha, \beta, \lambda) - \frac{1}{n+1}\right|}$$

$$\left[\Delta_3\left(x_{i:n} \mid \alpha, \beta, \lambda\right) - \Delta_3\left(x_{i-1:n} \mid \alpha, \beta, \lambda\right)\right] = 0,$$

where $D_i(\alpha, \beta, \lambda) \neq \frac{1}{n+1}$.

The estimators $\hat{\alpha}_{MSALDE}$, $\hat{\beta}_{MSALDE}$ and $\hat{\lambda}_{MSALDE}$ of α, β and λ can be obtained by solving the nonlinear equations

$$\frac{\partial}{\partial \alpha} T(\alpha, \beta, \lambda) = \sum_{i=1}^{n+1} \frac{\log D_i(\alpha, \beta, \lambda) - \log \frac{1}{n+1}}{\left|\log D_i(\alpha, \beta, \lambda) - \log \frac{1}{n+1}\right|} \frac{1}{D_i(\alpha, \beta, \lambda)}$$

$$[\Delta_1 (x_{i:n} \mid \alpha,\beta,\lambda) - \Delta_1 (x_{i-1:n} \mid \alpha,\beta,\lambda)] = 0$$

$$\frac{\partial}{\partial\beta}T(\alpha,\beta,\lambda) = \sum_{i=1}^{n+1} \frac{\log D_i(\alpha,\beta,\lambda) - \log\frac{1}{n+1}}{\left|\log D_i(\alpha,\beta,\lambda) - \log\frac{1}{n+1}\right|} \frac{1}{D_i(\alpha,\beta,\lambda)}$$

$$[\Delta_2 (x_{i:n} \mid \alpha,\beta,\lambda) - \Delta_2 (x_{i-1:n} \mid \alpha,\beta,\lambda)] = 0,$$

$$\frac{\partial}{\partial\lambda}T(\alpha,\beta,\lambda) = \sum_{i=1}^{n+1} \frac{\log D_i(\alpha,\beta,\lambda) - \log\frac{1}{n+1}}{\left|\log D_i(\alpha,\beta,\lambda) - \log\frac{1}{n+1}\right|} \frac{1}{D_i(\alpha,\beta,\lambda)}$$

$$[\Delta_3 (x_{i:n} \mid \alpha,\beta,\lambda) - \Delta_3 (x_{i-1:n} \mid \alpha,\beta,\lambda)] = 0,$$

where $\log D_i(\alpha,\beta,\lambda) \neq \log\frac{1}{n+1}$.

3.3 Methods of Ordinary and Weighted Least Squares

The least squares and weighted least squares estimators were proposed by Swain et al. (1988) to estimate the parameters of beta distribution (Swain et al., 1988). It is well known that

$$E\left[F (x_{i:n} \mid \alpha,\beta,\lambda)\right] = \frac{i}{n+1} \quad \text{and}$$
$$V\left[F(x_{i:n} \mid \alpha,\beta,\lambda)\right] = \frac{i\,(n-i+1)}{(n+1)^2\,(n+2)}. \tag{23}$$

Using the same notations as subsection 3.2, the ordinary least squares estimators $\widehat{\alpha}_{OLSE}$, $\widehat{\beta}_{OLSE}$ and $\widehat{\lambda}_{OLSE}$ of the parameters α, β and λ are obtained by minimizing the function:

$$S(\alpha,\beta,\lambda) = \sum_{i=1}^{n} \left[F (x_{i:n} \mid \alpha,\beta,\lambda) - \frac{i}{n+1}\right]^2. \tag{24}$$

These estimators can also be obtained by solving the following non-linear equations:

$$\sum_{i=1}^{n} \left[F (x_{i:n} \mid \alpha,\beta,\lambda) - \frac{i}{n+1}\right] \Delta_1 (x_{i:n} \mid \alpha,\beta,\lambda) = 0,$$

$$\sum_{i=1}^{n}\left[F\left(x_{i:n}\mid\alpha,\beta,\lambda\right)-\frac{i}{n+1}\right]\Delta_2\left(x_{i:n}\mid\alpha,\beta,\lambda\right)=0,$$
$$\sum_{i=1}^{n}\left[F\left(x_{i:n}\mid\alpha,\beta,\lambda\right)-\frac{i}{n+1}\right]\Delta_3\left(x_{i:n}\mid\alpha,\beta,\lambda\right)=0,$$

where $\Delta_1(.|\alpha,\beta,\lambda)$, $\Delta_2(.|\alpha,\beta,\lambda)$ and $\Delta_3(.|\alpha,\beta,\lambda)$ are given by Equations (17), (18) and (19), respectively.

The weighted least-squares estimators $\widehat{\alpha}_{WLSE}$, $\widehat{\beta}_{WLSE}$ and $\widehat{\lambda}_{WLSE}$ of the parameters α, β and λ are obtained by minimizing the function:

$$W\left(\alpha,\beta,\lambda\right)=\sum_{i=1}^{n}\frac{\left(n+1\right)^2\left(n+2\right)}{i\left(n-i+1\right)}\left[F\left(x_{i:n}\mid\alpha,\beta,\lambda\right)-\frac{i}{n+1}\right]^2. \quad (25)$$

The WLSE can be obtained by solving the following non-linear equations:

$$\sum_{i=1}^{n}\frac{\left(n+1\right)^2\left(n+2\right)}{i\left(n-i+1\right)}\left[F\left(x_{i:n}\mid\alpha,\beta,\lambda\right)-\frac{i}{n+1}\right]\Delta_1\left(x_{i:n}\mid\alpha,\beta,\lambda\right)=0, \quad (26)$$

$$\sum_{i=1}^{n}\frac{\left(n+1\right)^2\left(n+2\right)}{i\left(n-i+1\right)}\left[F\left(x_{i:n}\mid\alpha,\beta,\lambda\right)-\frac{i}{n+1}\right]\Delta_2\left(x_{i:n}\mid\alpha,\beta,\lambda\right)=0, \quad (27)$$

$$\sum_{i=1}^{n}\frac{\left(n+1\right)^2\left(n+2\right)}{i\left(n-i+1\right)}\left[F\left(x_{i:n}\mid\alpha,\beta,\lambda\right)-\frac{i}{n+1}\right]\Delta_3\left(x_{i:n}\mid\alpha,\beta,\lambda\right)=0, \quad (28)$$

3.4 Method of Percentiles

Since the PNH distribution has an explicit distribution function, the unknown parameters α, β and λ can be estimated by equating the sample percentile points with the population percentile points. This method is known as the percentile method (Kao, 1958, 1959). If p_i denote the estimate of $F(x_{i:n} \mid \alpha,\beta,\lambda)$, then the percentile estimators $\widehat{\alpha}_{PCE}\widehat{\beta}_{PCE}$ and $\widehat{\lambda}_{PCE}$ of the parameters α, β and λ can be obtained by minimizing the function $P(\alpha,\beta,\lambda)$ with respect to α, β and λ:

$$P(\alpha,\beta,\lambda)=\sum_{i=1}^{n}[x_{i:n}-F^{-1}(p_i|\alpha,\beta,\lambda)]^2,$$

where $p_i = \frac{i}{n+1}$ is the unbiased estimator of $F(x_{i:n} \mid \alpha, \beta, \lambda)$ and $F^{-1}(p_i|\alpha, \beta, \lambda)$ is defined in (8).

3.5 Methods of the Minimum Distances

This section considers three estimation methods by minimization of the goodness-of-fit statistics, i.e., minimizing the distance between the theoretical and empirical cumulative distribution functions, with respect to α, β and λ.

3.5.1 Method of Cramér-von-Mises

To motivate our choice of Cramér-von Mises type minimum distance estimators, MacDonald (1971) provided empirical evidence that the bias of the estimator is smaller than the other minimum distance estimators. Thus, the Cramér-von Mises estimators $\widehat{\alpha}_{CME}$, $\widehat{\beta}_{CME}$ and $\widehat{\lambda}_{CME}$ of the parameters α, β and λ are obtained by minimizing $C(\alpha, \beta, \lambda)$ with respect to α, β and λ:

$$C(\alpha, \beta, \lambda) = \frac{1}{12n} + \sum_{i=1}^{n} \left(F\left(x_{i:n} \mid \alpha, \beta, \lambda\right) - \frac{2i-1}{2n} \right)^2. \tag{29}$$

The estimators can be obtained by solving the following non-linear equations:

$$\sum_{i=1}^{n} \left(F\left(x_{i:n} \mid \alpha, \beta, \lambda\right) - \frac{2i-1}{2n} \right) \Delta_1\left(x_{i:n} \mid \alpha, \beta, \lambda\right) = 0,$$
$$\sum_{i=1}^{n} \left(F\left(x_{i:n} \mid \alpha, \beta, \lambda\right) - \frac{2i-1}{2n} \right) \Delta_2\left(x_{i:n} \mid \alpha, \beta, \lambda\right) = 0$$
$$\sum_{i=1}^{n} \left(F\left(x_{i:n} \mid \alpha, \beta, \lambda\right) - \frac{2i-1}{2n} \right) \Delta_3\left(x_{i:n} \mid \alpha, \beta, \lambda\right) = 0,$$

where $\Delta_1(. \mid \alpha, \beta, \lambda)$, $\Delta_2(. \mid \alpha, \beta, \lambda)$ and $\Delta_3(. \mid \alpha, \beta, \lambda)$ are given by (17), (18) and (19), respectively.

3.5.2 Methods of Anderson-Darling and Right-tail Anderson-Darling

Anderson and Darling **??** introduced a test as an alternative to statistical tests for detecting sample distributions departure from the normal distribution. This method is used here to obtain the Anderson-Darling estimators, $\widehat{\alpha}_{ADE}$, $\widehat{\beta}_{ADE}$ and $\widehat{\lambda}_{ADE}$ of the parameters α, β and λ, by minimizing the function

$A(\alpha,\beta,\lambda)$ with respect to α, β and λ respectively.

$$A(\alpha,\beta,\lambda) = -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)$$
$$\left\{\log F\left(x_{i:n} \mid \alpha,\beta,\lambda\right) + \log \overline{F}\left(x_{n+1-i:n} \mid \alpha,\beta,\lambda\right)\right\}. \quad (30)$$

The estimators can be obtained by solving the following non-linear equations:

$$\sum_{i=1}^{n}(2i-1)\left[\frac{\Delta_1\left(x_{i:n} \mid \alpha,\beta,\lambda\right)}{F\left(x_{i:n} \mid \alpha,\beta,\sigma\right)} - \frac{\Delta_1\left(x_{n+1-i:n} \mid \alpha,\beta,\lambda\right)}{\overline{F}\left(x_{n+1-i:n} \mid \alpha,\beta,\lambda\right)}\right] = 0,$$
$$\sum_{i=1}^{n}(2i-1)\left[\frac{\Delta_2\left(x_{i:n} \mid c,\beta,\theta\right)}{F\left(x_{i:n} \mid \alpha,\beta,\lambda\right)} - \frac{\Delta_2\left(x_{n+1-i:n} \mid \alpha,\beta,\lambda\right)}{\overline{F}\left(x_{n+1-i:n} \mid \alpha,\beta,\lambda\right)}\right] = 0,$$
$$\sum_{i=1}^{n}(2i-1)\left[\frac{\Delta_3\left(x_{i:n} \mid \alpha,\beta,\lambda\right)}{F\left(x_{i:n} \mid c,\beta,\theta\right)} - \frac{\Delta_3\left(x_{n+1-i:n} \mid \alpha,\beta,\lambda\right)}{\overline{F}\left(x_{n+1-i:n} \mid \alpha,\beta,\lambda\right)}\right] = 0.$$

The right-tail Anderson-Darling estimators $\widehat{\alpha}_{RTADE}$, $\widehat{\beta}_{RTADE}$ and $\widehat{\lambda}_{RTADE}$ of the parameters α, β and λ are obtained by minimizing $R(\alpha,\beta,\lambda)$ with respect to α, β and λ. The right-tail Anderson-Darling is defined as

$$R(\alpha,\beta,\lambda) = \frac{n}{2} - 2\sum_{i=1}^{n} F\left(x_{i:n} \mid \alpha,\beta,\lambda\right)$$
$$-\frac{1}{n}\sum_{i=1}^{n}(2i-1)\log \overline{F}\left(x_{n+1-i:n} \mid \alpha,\beta,\lambda\right). \quad (31)$$

The estimators can also be obtained by solving the following non-linear equations.

$$-2\sum_{i=1}^{n}\frac{\Delta_1\left(x_{i:n} \mid \alpha,\beta,\lambda\right)}{F\left(x_{i:n} \mid \alpha,\beta,\lambda\right)} + \frac{1}{n}\sum_{i=1}^{n}(2i-1)\frac{\Delta_1\left(x_{n+1-i:n} \mid \alpha,\beta,\lambda\right)}{\overline{F}\left(x_{n+1-i:n} \mid \alpha,\beta,\lambda\right)} = 0,$$
$$-2\sum_{i=1}^{n}\frac{\Delta_2\left(x_{i:n} \mid \alpha,\beta,\lambda\right)}{F\left(x_{i:n} \mid \alpha,\beta,\lambda\right)} + \frac{1}{n}\sum_{i=1}^{n}(2i-1)\frac{\Delta_2\left(x_{n+1-i:n} \mid \alpha,\beta,\lambda\right)}{\overline{F}\left(x_{n+1-i:n} \mid \alpha,\beta,\lambda\right)} = 0,$$
$$-2\sum_{i=1}^{n}\frac{\Delta_3\left(x_{i:n} \mid \alpha,\beta,\lambda\right)}{F\left(x_{i:n} \mid \alpha,\beta,\lambda\right)} + \frac{1}{n}\sum_{i=1}^{n}(2i-1)\frac{\Delta_3\left(x_{n+1-i:n} \mid \alpha,\beta,\lambda\right)}{\overline{F}\left(x_{n+1-i:n} \mid \alpha,\beta,\lambda\right)} = 0,$$

where $\Delta_1(\cdot \mid \alpha, \beta, \lambda)$, $\Delta_2(\cdot \mid \alpha, \beta, \lambda)$ and $\Delta_3(\cdot \mid \alpha, \beta, \lambda)$ are given by Equations (17), (18) and (19), respectively.

4 Simulation Study

This section discusses Monte Carlo simulation studies to assess the performance of the frequentist estimators mentioned in the previous section. In particular, we used bias, root mean squared error, the average absolute difference between the theoretical and the empirical estimate of the distribution functions, and the maximum absolute difference between the theoretical and the empirical distribution functions as the performance assessment criteria. For comparison, we considered the following sample sizes: n = 20, 40, 60, 80, 100. Ten thousand independent samples of the aforementioned sample sizes were generated from PNH distribution with parameters $(\alpha, \beta, \lambda) = \{(0.5, 0.5, 0.5), (3.5, 3.5, 3.5)\}$. It is observed that 10,000 repetitions are sufficiently large to have stable results. For all the methods considered in this study, first we have estimated the parameters using the method of maximum likelihood and used them as the initial values for the rest of the methods. Also, the same randomly generated samples are used to compute the simulation summaries of different estimation methods. The results of the simulation studies are tabulated in Tables 1–2.

For each estimate we calculate the bias, root mean-squares error (RMSE), the average absolute difference between the theoretical and the empirical estimate of the distribution functions (D_{abs}), and the maximum absolute difference between the theoretical and the empirical distribution functions (D_{max}). The statistics are obtained using the following formulae:

$$\text{Bias}(\hat{\alpha}) = \frac{1}{R}\sum_{i=1}^{R}(\hat{\alpha}_i - \alpha), \quad \text{Bias}(\hat{\lambda}) = \frac{1}{R}\sum_{i=1}^{R}(\hat{\lambda}_i - \lambda) \tag{32}$$

$$\text{RMSE}(\hat{\alpha}) = \sqrt{\frac{1}{R}\sum_{i=1}^{R}(\hat{\alpha}_i - \alpha)^2}, \quad \text{RMSE}(\hat{\lambda}) = \sqrt{\frac{1}{R}\sum_{i=1}^{R}(\hat{\lambda}_i - \lambda)^2} \tag{33}$$

$$D_{\text{abs}}(\hat{\alpha}) = \frac{1}{(nR)}\sum_{i=1}^{R}\sum_{j=1}^{n}\left|F(x_{ij}|\alpha, \lambda) - F(x_{ij}|\hat{\alpha}, \hat{\lambda})\right| \tag{34}$$

$$D_{\max}(\hat{\alpha}) = \frac{1}{R}\sum_{i=1}^{R}\max_{j}\left|F(x_{ij}|\alpha, \lambda) - F(x_{ij}|\hat{\alpha}, \hat{\lambda})\right| \tag{35}$$

Simulated bias, RMSE, D_{abs}, $D_{\max}$ for the estimates are listed in Tables 1–2. The row with label $\sum$ Ranks shows the partial sum of the ranks and superscript indicates the rank of each of the estimators among all the estimators for that metric. For example, Table 1 shows the bias of MLE($\hat{\alpha}$) as 0.414^4 for $n = 20$. This indicates, the bias of $\hat{\alpha}$ obtained using the method of maximum likelihood ranks 4th among all other estimators.

The following observations can be drawn from the Tables 1–2.

1. All the estimators show the property of consistency, i.e., the RMSE decreased as sample size increased.
2. The bias of $\hat{\alpha}$ decreased by increasing n for all the method of estimations.
3. The bias of $\hat{\beta}$ decreased by increasing n for all the method of estimations.
4. The bias of $\hat{\lambda}$ decreased by increasing n for all the method of estimations but for small sample size, the estimate of λ is highly biased.
5. The bias of MSALDE increased by increasing n as compared to the other methods.
6. D_{abs} is smaller than $D_{\max}$ for all the estimation techniques. Again, the statistics get smaller with the increase of sample size.
7. In terms of performance of the methods of estimation, it is observed that the WLS and MPS estimators uniformly produce the least biases of the estimates with least RMSE for most of the configurations considered in our studies.

5 Bayesian Estimation

This section presents the Bayesian inference of the unknown parameters of the PHN distribution. It is needless to mention that, if all the parameters of the model are unknown, a joint conjugate prior for the parameters does not exist. For this, we assume piecewise independent priors and the proposed priors for the parameters α, β and λ may be taken as $\alpha \sim Gamma(a, b), \beta \sim Gamma(c, d)$ and $\lambda \sim Gamma(e, f)$. The joint prior distribution of α, β and λ can be written as $p(\alpha, \beta, \lambda) \propto \alpha^{a-1}\beta^{c-1}\lambda^{e-1}\exp(-b\alpha - d\beta - f\lambda)$. We assume $\boldsymbol{\theta} = (\alpha, \beta, \lambda), \boldsymbol{x} = (x_1, x_2, \cdots, x_n)$, $P(\boldsymbol{\theta})$ denote the joint posterior and $L(\boldsymbol{\theta}; \boldsymbol{x})$ is the likelihood function. For the $PNH(x|\alpha, \beta, \lambda)$,

Table 1 Simulation results for $\alpha = \beta = \lambda = 0.5$

n	Est.	MLE	LSE	WLS	PCE	MPS	MSADE	MSALDE	CVM	AD	RAD
20	Bias($\hat{\alpha}$)	0.414^{4}	0.228^{2}	1.314^{6}	43.300^{10}	0.174^{1}	1.641^{8}	5.560^{9}	0.338^{3}	1.381^{7}	1.186^{5}
	RMSE($\hat{\alpha}$)	0.864^{3}	0.751^{2}	3.416^{6}	99.946^{10}	0.578^{1}	3.610^{7}	8.457^{9}	0.901^{4}	3.899^{8}	3.098^{5}
	Bias($\hat{\beta}$)	0.408^{7}	0.223^{5}	0.147^{2}	30.437^{10}	0.047^{1}	2.762^{8}	6.531^{9}	0.212^{4}	0.172^{3}	0.320^{6}
	RMSE($\hat{\beta}$)	2.520^{7}	0.872^{5}	0.548^{2}	56.548^{10}	0.314^{1}	5.002^{8}	9.736^{9}	0.554^{3}	0.596^{4}	0.968^{6}
	Bias($\hat{\lambda}$)	135.204^{10}	15.598^{9}	6.343^{6}	0.010^{1}	1.580^{2}	1.586^{3}	4.137^{5}	3.009^{4}	7.300^{7}	8.762^{8}
	RMSE($\hat{\lambda}$)	1884.978^{10}	132.385^{9}	36.852^{6}	0.010^{1}	4.980^{3}	4.166^{2}	7.397^{5}	7.098^{4}	40.617^{7}	43.115^{8}
	D_{abs}	$0.059^{1.5}$	0.062^{5}	0.061^{3}	0.182^{7}	$0.059^{1.5}$	0.421^{10}	0.399^{9}	0.062^{5}	0.310^{8}	0.062^{5}
	$D_{\max}$	0.102^{2}	0.106^{4}	0.104^{3}	0.299^{7}	0.098^{1}	0.819^{10}	0.788^{9}	$0.109^{5.5}$	0.547^{8}	$0.109^{5.5}$
	$\sum$Ranks	44.5^{5}	41^{4}	34^{3}	$56^{8.5}$	11.5^{1}	$56^{8.5}$	64^{10}	32.5^{2}	52^{7}	48.5^{6}
40	Bias($\hat{\alpha}$)	0.217^{3}	0.174^{2}	0.612^{7}	15.429^{10}	0.088^{1}	2.108^{8}	5.988^{9}	0.221^{4}	0.421^{6}	0.410^{5}
	RMSE($\hat{\alpha}$)	0.494^{2}	0.523^{3}	2.065^{7}	43.918^{10}	0.360^{1}	4.475^{8}	9.053^{9}	0.584^{4}	1.760^{6}	1.574^{5}
	Bias($\hat{\beta}$)	0.088^{5}	0.079^{4}	0.051^{2}	18.925^{10}	0.008^{1}	3.061^{8}	6.758^{9}	0.093^{6}	0.058^{3}	0.116^{7}
	RMSE($\hat{\beta}$)	0.518^{7}	0.342^{5}	0.247^{3}	51.891^{10}	0.161^{1}	5.661^{8}	10.056^{9}	0.273^{4}	0.237^{2}	0.411^{6}
	Bias($\hat{\lambda}$)	9.389^{10}	4.579^{9}	1.662^{5}	0.010^{1}	0.536^{2}	1.884^{7}	4.098^{8}	1.322^{3}	1.478^{4}	1.855^{6}
	RMSE($\hat{\lambda}$)	196.246^{10}	47.051^{9}	17.031^{8}	0.010^{1}	2.404^{2}	4.377^{4}	7.285^{5}	4.194^{3}	13.045^{7}	10.245^{6}
	D_{abs}	$0.042^{1.5}$	0.044^{5}	0.043^{3}	0.129^{7}	$0.042^{1.5}$	0.437^{10}	0.404^{9}	0.044^{5}	0.246^{8}	0.044^{5}
	$D_{\max}$	0.073^{2}	0.077^{4}	0.075^{3}	0.213^{7}	0.071^{1}	0.851^{10}	0.794^{9}	0.079^{6}	0.408^{8}	0.078^{5}
	$\sum$Ranks	40.5^{4}	41^{5}	38^{3}	56^{8}	10.5^{1}	63^{9}	67^{10}	35^{2}	44^{6}	45^{7}
60	Bias($\hat{\alpha}$)	0.148^{3}	0.146^{2}	0.322^{7}	14.904^{10}	0.055^{1}	2.497^{8}	6.045^{9}	0.174^{4}	0.191^{5}	0.199^{6}
	RMSE($\hat{\alpha}$)	0.366^{2}	0.438^{3}	1.260^{7}	54.299^{10}	0.275^{1}	5.151^{8}	9.186^{9}	0.473^{4}	0.911^{6}	0.897^{5}
	Bias($\hat{\beta}$)	0.038^{4}	0.042^{5}	0.026^{2}	18.419^{10}	-0.001^{1}	3.248^{8}	6.832^{9}	0.055^{6}	0.032^{3}	0.061^{7}
	RMSE($\hat{\beta}$)	0.233^{7}	0.216^{5}	0.156^{3}	60.180^{10}	0.113^{1}	6.011^{8}	10.120^{9}	0.187^{4}	0.143^{2}	0.229^{6}
	Bias($\hat{\lambda}$)	2.411^{9}	1.559^{7}	0.577^{4}	0.010^{1}	0.272^{2}	2.084^{8}	4.152^{10}	0.715^{6}	0.491^{3}	0.707^{5}
	RMSE($\hat{\lambda}$)	49.611^{10}	20.585^{9}	7.194^{7}	0.010^{1}	1.386^{2}	4.606^{6}	7.401^{8}	2.736^{3}	4.208^{4}	4.238^{5}
	D_{abs}	0.034^{1}	0.036^{5}	$0.035^{2.5}$	0.117^{7}	$0.035^{2.5}$	0.443^{10}	0.406^{9}	0.036^{5}	0.212^{8}	0.036^{5}
	$D_{\max}$	0.059^{2}	$0.064^{4.5}$	0.061^{3}	0.192^{7}	0.058^{1}	0.861^{10}	0.797^{9}	0.065^{6}	0.340^{8}	$0.064^{4.5}$
	$\sum$Ranks	$38^{3.5}$	40.5^{6}	35.5^{2}	56^{8}	11.5^{1}	66^{9}	72^{10}	$38^{3.5}$	39^{5}	43.5^{7}

80	Bias($\hat{\alpha}$)	0.109^{2}	0.124^{5}	0.198^{7}	9.327^{10}	0.036^{1}	2.663^{8}	6.092^{9}	0.144^{6}	0.110^{3}	0.121^{4}
	RMSE($\hat{\alpha}$)	0.292^{2}	0.385^{3}	0.893^{7}	46.999^{10}	0.225^{1}	5.445^{8}	9.241^{9}	0.409^{4}	0.567^{5}	0.591^{6}
	Bias($\hat{\beta}$)	0.021^{3}	0.027^{5}	0.017^{2}	15.827^{10}	-0.003^{1}	3.292^{8}	6.981^{9}	0.038^{6}	0.022^{4}	0.041^{7}
	RMSE($\hat{\beta}$)	$0.143^{4.5}$	0.150^{6}	0.112^{3}	61.087^{10}	0.090^{1}	6.079^{8}	10.321^{9}	$0.143^{4.5}$	0.108^{2}	0.157^{7}
	Bias($\hat{\lambda}$)	0.680^{7}	0.696^{8}	$0.263^{3.5}$	0.010^{1}	0.171^{2}	2.291^{9}	4.283^{10}	0.444^{6}	$0.263^{3.5}$	0.391^{5}
	RMSE($\hat{\lambda}$)	18.605^{10}	10.232^{9}	1.730^{3}	0.010^{1}	0.856^{2}	4.875^{7}	7.566^{8}	1.863^{4}	1.970^{5}	2.054^{6}
	D_{abs}	0.030^{2}	$0.032^{5.5}$	0.030^{2}	0.101^{7}	0.030^{2}	0.449^{10}	0.405^{9}	$0.032^{5.5}$	0.191^{8}	0.031^{4}
	$D_{\max}$	$0.051^{1.5}$	$0.055^{4.5}$	0.053^{3}	0.163^{7}	$0.051^{1.5}$	0.869^{10}	0.791^{9}	0.056^{6}	0.298^{8}	$0.055^{4.5}$
	$\sum$Ranks	32^{3}	46^{7}	30.5^{2}	56^{8}	11.5^{1}	68^{9}	72^{10}	42^{5}	38.5^{4}	43.5^{6}
100	Bias($\hat{\alpha}$)	0.085^{4}	0.103^{5}	0.132^{7}	4.508^{9}	0.025^{1}	2.772^{8}	5.790^{10}	0.117^{6}	0.070^{2}	0.077^{3}
	RMSE($\hat{\alpha}$)	0.243^{2}	0.334^{3}	0.670^{7}	23.179^{10}	0.190^{1}	5.631^{8}	8.993^{9}	0.350^{5}	0.341^{4}	0.383^{6}
	Bias($\hat{\beta}$)	0.014^{3}	0.018^{5}	0.012^{2}	13.326^{10}	-0.003^{1}	3.326^{8}	7.021^{9}	0.028^{6}	0.016^{4}	0.030^{7}
	RMSE($\hat{\beta}$)	0.103^{4}	0.110^{5}	$0.089^{2.5}$	60.285^{10}	0.078^{1}	6.125^{8}	10.333^{9}	0.115^{6}	$0.089^{2.5}$	0.121^{7}
	Bias($\hat{\lambda}$)	0.262^{6}	0.287^{7}	0.161^{3}	0.010^{1}	0.124^{2}	2.396^{9}	4.095^{10}	0.289^{8}	0.170^{4}	0.247^{5}
	RMSE($\hat{\lambda}$)	8.255^{10}	2.617^{7}	0.656^{3}	0.010^{1}	0.599^{2}	5.021^{8}	7.337^{9}	1.199^{6}	0.674^{4}	1.011^{5}
	D_{abs}	0.027^{2}	0.028^{5}	0.027^{2}	0.088^{7}	0.027^{2}	0.452^{10}	0.406^{9}	0.028^{5}	0.173^{8}	0.028^{5}
	$D_{\max}$	0.046^{2}	$0.049^{4.5}$	0.047^{3}	0.141^{7}	0.045^{1}	0.874^{10}	0.788^{9}	0.050^{6}	0.266^{8}	$0.049^{4.5}$
	$\sum$Ranks	33^{3}	41.5^{5}	29.5^{2}	55^{8}	11^{1}	69^{9}	74^{10}	48^{7}	36.5^{4}	42.5^{6}

Table 2 Simulation results for $\alpha = \beta = \lambda = 3.5$

n	Est.	MLE	LSE	WLS	PCE	MPS	MSADE	MSALDE	CVM	AD	RAD
20	Bias($\hat{\alpha}$)	21.500^{10}	3.754^{7}	1.620^{4}	3.415^{6}	5.791^{9}	1.055^{2}	1.053^{1}	4.066^{8}	2.131^{5}	1.505^{3}
	RMSE($\hat{\alpha}$)	36.500^{10}	7.952^{8}	4.533^{3}	4.880^{5}	10.509^{9}	4.814^{4}	4.146^{1}	7.898^{7}	5.166^{6}	4.360^{2}
	Bias($\hat{\beta}$)	3.500^{10}	2.261^{8}	1.048^{2}	-3.498^{9}	0.451^{1}	2.017^{6}	1.913^{5}	1.587^{4}	1.406^{3}	2.165^{7}
	RMSE($\hat{\beta}$)	18.500^{10}	5.733^{8}	3.045^{2}	4.131^{5}	2.135^{1}	5.680^{7}	4.210^{6}	3.353^{3}	3.467^{4}	6.983^{9}
	Bias($\hat{\lambda}$)	14.500^{9}	16.193^{10}	6.280^{6}	0.010^{1}	4.287^{4}	0.517^{2}	0.759^{3}	4.620^{5}	6.839^{7}	7.728^{8}
	RMSE($\hat{\lambda}$)	116.500^{10}	39.933^{9}	15.753^{6}	0.010^{1}	9.311^{4}	4.007^{3}	3.759^{2}	9.548^{5}	19.058^{7}	20.687^{8}
	D_{abs}	0.049^{1}	0.062^{6}	0.059^{3}	0.512^{10}	0.058^{2}	0.467^{9}	0.455^{8}	$0.061^{4.5}$	0.410^{7}	$0.061^{4.5}$
	$D_{\max}$	0.85^{7}	0.101^{3}	0.097^{2}	0.961^{10}	0.093^{1}	0.925^{9}	0.908^{8}	0.103^{5}	0.750^{6}	0.102^{4}
	$\sum$Ranks	67^{10}	59^{9}	28^{1}	47^{8}	31^{2}	42^{5}	34^{3}	41.5^{4}	45^{6}	45.5^{7}
40	Bias($\hat{\alpha}$)	18.105^{10}	3.097^{6}	1.413^{4}	4.912^{9}	4.735^{8}	1.141^{3}	0.603^{1}	3.333^{7}	1.614^{5}	1.120^{2}
	RMSE($\hat{\alpha}$)	32.645^{10}	6.786^{7}	4.027^{3}	5.886^{6}	9.455^{9}	4.778^{5}	3.682^{2}	6.849^{8}	4.260^{4}	3.636^{1}
	Bias($\hat{\beta}$)	2.531^{9}	1.380^{6}	0.688^{2}	-3.370^{10}	0.327^{1}	1.391^{7}	1.884^{8}	0.998^{4}	0.781^{3}	1.133^{5}
	RMSE($\hat{\beta}$)	16.122^{10}	3.351^{6}	1.821^{2}	3.874^{10}	1.552^{1}	5.193^{7}	4.281^{8}	2.212^{4}	1.953^{3}	2.427^{5}
	Bias($\hat{\lambda}$)	13.578^{10}	9.325^{9}	4.098^{7}	0.010^{1}	3.317^{4}	0.831^{3}	0.386^{2}	4.020^{6}	3.970^{5}	4.492^{8}
	RMSE($\hat{\lambda}$)	108.529^{10}	23.107^{9}	10.382^{7}	0.010^{1}	7.900^{4}	4.296^{3}	3.438^{2}	8.812^{5}	10.137^{6}	10.624^{8}
	D_{abs}	$0.041^{1.5}$	0.044^{6}	0.042^{3}	0.512^{10}	$0.041^{1.5}$	0.475^{9}	0.465^{8}	$0.043^{4.5}$	0.396^{7}	$0.043^{4.5}$
	$D_{\max}$	0.070^{3}	0.073^{5}	0.069^{2}	0.983^{10}	0.067^{1}	0.956^{9}	0.929^{8}	0.073^{5}	0.708^{7}	0.073^{5}
	$\sum$Ranks	63.5^{10}	$54^{8.5}$	30^{2}	$54^{8.5}$	29.5^{1}	48^{7}	39^{4}	43.5^{6}	40^{5}	38.5^{3}
60	Bias($\hat{\alpha}$)	14.247^{10}	2.712^{6}	1.285^{3}	5.135^{9}	4.161^{8}	1.534^{5}	0.704^{1}	2.896^{7}	1.400^{4}	0.974^{2}
	RMSE($\hat{\alpha}$)	27.222^{10}	6.081^{6}	3.698^{3}	6.196^{8}	8.785^{9}	4.961^{5}	3.627^{2}	6.165^{7}	3.845^{4}	3.277^{1}
	Bias($\hat{\beta}$)	0.956^{6}	1.009^{7}	0.523^{2}	-3.380^{10}	0.246^{1}	1.046^{8}	1.710^{9}	0.779^{4}	0.563^{3}	0.809^{5}
	RMSE($\hat{\beta}$)	5.850^{10}	2.592^{6}	1.464^{2}	3.864^{7}	1.279^{1}	4.954^{9}	4.114^{8}	1.810^{5}	1.506^{3}	1.807^{4}
	Bias($\hat{\lambda}$)	4.585^{9}	6.574^{10}	3.050^{6}	0.010^{1}	2.629^{4}	1.185^{3}	0.482^{2}	3.522^{8}	2.923^{5}	3.310^{7}
	RMSE($\hat{\lambda}$)	34.637^{10}	16.481^{9}	7.975^{6}	0.010^{1}	6.876^{4}	4.426^{3}	3.406^{2}	8.166^{8}	7.842^{5}	8.155^{7}
	D_{abs}	0.034^{2}	$0.035^{4.5}$	0.034^{2}	0.505^{10}	0.034^{2}	0.477^{9}	0.468^{8}	0.036^{6}	0.384^{7}	$0.035^{4.5}$
	$D_{\max}$	0.056^{2}	0.060^{5}	0.057^{3}	0.986^{10}	0.055^{1}	0.965^{9}	0.932^{8}	0.060^{5}	0.674^{7}	0.060^{5}
	$\sum$Ranks	59^{10}	53.5^{8}	27^{1}	56^{9}	30^{2}	51^{7}	40^{5}	50^{6}	38^{4}	35.5^{3}

80	Bias($\hat{\alpha}$)	11.557^{10}	2.427^{6}	1.158^{3}	5.976^{9}	3.698^{8}	1.762^{5}	0.808^{1}	2.573^{7}	1.221^{4}	0.879^{2}
	RMSE($\hat{\alpha}$)	23.196^{10}	5.605^{6}	3.452^{2}	7.198^{8}	8.218^{9}	5.111^{5}	3.637^{4}	5.684^{7}	3.511^{3}	3.084^{1}
	Bias($\hat{\beta}$)	0.538^{4}	0.814^{7}	0.434^{2}	-3.449^{10}	0.199^{1}	0.867^{8}	1.648^{9}	0.673^{6}	0.461^{3}	0.657^{5}
	RMSE($\hat{\beta}$)	2.768^{7}	2.155^{6}	1.252^{2}	3.917^{8}	1.129^{1}	4.934^{10}	4.085^{9}	1.613^{5}	1.295^{3}	1.527^{4}
	Bias($\hat{\lambda}$)	2.375^{5}	5.245^{10}	2.502^{7}	0.010^{1}	2.162^{4}	1.479^{3}	0.632^{2}	3.250^{9}	2.403^{6}	2.725^{8}
	RMSE($\hat{\lambda}$)	14.973^{10}	13.319^{9}	6.661^{6}	0.010^{1}	6.074^{4}	4.648^{3}	3.498^{2}	7.757^{8}	6.582^{5}	6.782^{7}
	D_{abs}	$0.029^{1.5}$	0.031^{5}	0.030^{3}	0.503^{10}	$0.029^{1.5}$	0.479^{9}	0.470^{8}	0.031^{5}	0.373^{7}	0.031^{5}
	$D_{\max}$	$0.049^{2.5}$	0.052^{5}	$0.049^{2.5}$	0.991^{10}	0.048^{1}	0.970^{9}	0.932^{8}	0.052^{5}	0.646^{7}	0.052^{5}
	$\sum$Ranks	50^{6}	54^{9}	27.5^{1}	57^{10}	29.5^{2}	$52^{7.5}$	43^{5}	$52^{7.5}$	38^{4}	37^{3}
100	Bias($\hat{\alpha}$)	9.765^{10}	2.213^{6}	1.074^{3}	7.395^{9}	3.358^{8}	1.923^{5}	0.909^{2}	2.332^{7}	1.111^{4}	0.818^{1}
	RMSE($\hat{\alpha}$)	20.457^{10}	5.205^{5}	3.222^{2}	8.337^{9}	7.779^{8}	5.296^{7}	3.634^{4}	5.268^{6}	3.255^{3}	2.903^{1}
	Bias($\hat{\beta}$)	0.361^{2}	0.666^{7}	0.363^{3}	-3.374^{10}	0.166^{1}	0.772^{8}	1.688^{9}	0.585^{6}	0.377^{4}	0.548^{5}
	RMSE($\hat{\beta}$)	1.505^{6}	1.842^{7}	1.112^{2}	3.741^{8}	1.018^{1}	4.805^{10}	4.090^{9}	1.471^{5}	1.117^{3}	1.324^{4}
	Bias($\hat{\lambda}$)	1.509^{3}	4.234^{10}	2.083^{7}	0.010^{1}	1.839^{5}	1.639^{4}	0.764^{2}	2.915^{9}	1.988^{6}	2.307^{8}
	RMSE($\hat{\lambda}$)	7.579^{9}	11.127^{10}	5.823^{6}	0.010^{1}	5.488^{4}	4.793^{3}	3.562^{2}	7.285^{8}	5.740^{5}	5.931^{7}
	D_{abs}	0.026^{2}	$0.028^{5.5}$	0.026^{2}	0.502^{10}	0.026^{2}	0.481^{9}	0.472^{8}	$0.028^{5.5}$	0.362^{7}	0.027^{4}
	$D_{\max}$	$0.043^{1.5}$	$0.047^{5.5}$	0.044^{3}	0.992^{10}	$0.043^{1.5}$	0.973^{9}	0.934^{8}	$0.047^{5.5}$	0.620^{7}	0.046^{4}
	$\sum$Ranks	43.5^{5}	56^{9}	28^{1}	58^{10}	30.5^{2}	55^{8}	44^{6}	52^{7}	39^{4}	34^{3}

the likelihood function can be written as

$$L(\boldsymbol{\theta};\boldsymbol{x}) \propto (\alpha\beta\lambda)^n \prod_{i=1}^{n}(1+\lambda x_i)^{\alpha-1} \exp\biggl(-\sum_{i=1}^{n}(1+\lambda x_i)^{\alpha}\biggr)$$
$$\prod_{i=1}^{n}\biggl\{1-\exp\biggl(1-(1+\lambda x_i)^{\alpha}\biggr)\biggr\}^{\beta-1}$$
$$\times \exp\biggl(-\sum_{i=1}^{n}\biggl\{1-\exp\biggl(1-(1+\lambda x_i)^{\alpha}\biggr)\biggr\}^{\beta}\biggr) \tag{36}$$

Therefore, we write the joint posterior as

$$P(\alpha,\beta,\lambda|\boldsymbol{x}) \propto \alpha^{n+a-1}\beta^{n+c-1}\lambda^{n+e-1}$$
$$\exp\biggl(-\alpha\biggl\{b-\sum_{i=1}^{n}\ln(1+\lambda x_i)\biggr\}\biggr)$$
$$\exp\biggl(-\sum_{i=1}^{n}\ln(1+\lambda x_i)\biggr)$$
$$\exp\biggl(-\beta\biggl\{d-\sum_{i=1}^{n}\ln\biggl[1-\exp\biggl(1-(1+\lambda x_i)^{\alpha}\biggr)\biggr]\biggr\}\biggr)$$
$$\exp\biggl(-\sum_{i=1}^{n}\ln\biggl\{1-\exp\biggl(1-(1+\lambda x_i)^{\alpha}\biggr)\biggr\}\biggr)\exp(-\lambda f)$$
$$\exp\biggl(-\sum_{i=1}^{n}\biggl\{1-\exp\biggl(1-(1+\lambda x_i)^{\alpha}\biggr)\biggr\}^{\beta}\biggr) \tag{37}$$

$$P(\alpha,\beta,\lambda|\boldsymbol{x}) \propto P_{\alpha}\biggl(n+a, b-\sum_{i=1}^{n}\ln(1+\lambda x_i)|\boldsymbol{x},\lambda\biggr)P(\lambda|\boldsymbol{x}\alpha,\beta)$$
$$P_{\beta}\biggl(n+c, d-\sum_{i=1}^{n}\ln\biggl[1-\exp\biggl(1-(1+\lambda x_i)^{\alpha}\biggr)\biggr]|\boldsymbol{x},\alpha,\lambda\biggr) \tag{38}$$

where P_α and P_β are the gamma densities, and $P(\lambda|\boldsymbol{x}\alpha,\beta) = \lambda^{n+e-1}\exp(-\sum_{i=1}^{n}\ln\{1-\exp(1-(1+\lambda x_i)^{\alpha})\})\exp(-\lambda f)\exp(-\sum_{i=1}^{n}\{1-\exp(1-(1+\lambda x_i)^{\alpha})\}^{\beta})\exp(-\sum_{i=1}^{n}\ln(1+\lambda x_i))$. It is not difficult to show that $P(\lambda|\boldsymbol{x}, \alpha, \beta)$ is log-concave for $\beta > 2$ and $\alpha < 1$ and thus, the idea of Devroye (1986) can be used. Here, we will implement the Metropolis Hastings (MH) (Metropolis et al., 1953) algorithm to compute the estimators. The MH algorithm is a powerful Markov Chain Monte Carlo algorithm. To this end, we assume gamma density as transition kernel $q(\lambda^{(i)}|\lambda^{(*)})$ for sampling value of λ. The choice of gamma distribution has been considered purely for illustration purpose, and other suitable distributions can be used. After generating the marginal densities, the next step is to calculate the posterior summaries, $\mathbb{E}(\boldsymbol{\theta}|\boldsymbol{x}) = \int_{\boldsymbol{\theta}} \boldsymbol{\theta}\mathbb{P}(\boldsymbol{\theta}|\boldsymbol{x})$. The steps to calculate the Bayes estimates are as follow:

Step 1: Take some initial guess values of α, β and λ, say α_0, β_0 and λ_0, respectively;

1. To generate λ, evaluate the acceptance probability by $k(\lambda^{(i)}, \lambda^{(*)}) = \min(1, \frac{P(\lambda^{(*)}|\boldsymbol{x})q(\lambda^{(i)}|\lambda^{(*)})}{P(\lambda^{(i)}|\boldsymbol{x})q(\lambda^{(*)}|\lambda^{(i)})})$, where $P(\lambda|\boldsymbol{x})$ has been defined above.
2. Generate a random numbers u from $Uniform(0, 1)$
3. If $k(\lambda^{(i)}, \lambda^{(*)}) \geq u$, $\lambda^{(i+1)} = \lambda^{(*)}$, otherwise $\lambda^{(i+1)} = \lambda^{(i)}$.

Step 2: Suppose at the ith step, α, β and λ take the values α_i, β_i and λ_i. Now we can generate $\mathbb{P}(\lambda_{i+1}|\alpha_i, \beta_i, \boldsymbol{x})$, $\mathbb{P}(\alpha_{i+1}|\lambda_i, \boldsymbol{x})$ and $\mathbb{P}(\beta_{i+1}|\alpha_i, \lambda_i, \boldsymbol{x})$;

Step 3: Repeat the above step N times;

Step 4: Calculate the Bayes estimator of $h(\alpha, \lambda)$ by $\frac{1}{N-M}\sum_{i=M+1}^{N} h(\alpha_i, \lambda_i)$, where M denote the number of burn-in sample.

For the Bayesian analysis, we generated $12,000$ samples for α, β and λ, and the Bayes estimates with other posterior summaries, like MCMC error, median, 95% Bayesian intervals have been tabulated in Table 3 for the above mentioned parameter combinations and sample sizes. To compute the posterior summaries, we selected the hyperparameters in such a way that mean of the priors equal to the nominal parameter values with large variances. Moreover, we have used $M = 2,000$ as the burn-in period for our calculations. From Table 3, it is noticed that as the sample size increases, the Bayes estimates approaches to the nominal values and the Bayesian intervals become tighter for large sample sizes. Furthermore, the MCMC error decreases with the increase of sample sizes.

Table 3 Monte Carlo Markov Chain results for Bayesian analysis

Parameter	n	Estimate	SD	MC error	95% CI	Median
$\alpha = 0.5$	20	0.5044	0.5101	0.0060	(0.0128,1.88)	0.3504
	40	0.5051	0.5085	0.0036	(0.0127,1.855)	0.3464
	60	0.4999	0.5005	0.0014	(0.0128,1.833)	0.3468
	80	0.501	0.5006	0.0012	(0.0129,1.824)	0.3464
	100	0.5001	0.5002	0.0004	(0.0127,1.804)	0.3468
$\beta = 0.5$	20	0.4974	0.5037	0.0048	(0.0123,1.919)	0.3402
	40	0.4997	0.5035	0.0033	(0.0119,1.893)	0.3447
	60	0.5008	0.5019	0.0014	(0.0128,1.856)	0.3467
	80	0.5011	0.5014	0.0010	(0.0132,1.849)	0.3475
	100	0.5001	0.5004	0.0004	(0.0127,1.846)	0.3469
$\lambda = 0.5$	20	0.4986	0.4993	0.0045	(0.0124,1.845)	0.3458
	40	0.503	0.5119	0.0031	(0.0116,1.897)	0.3489
	60	0.4996	0.4989	0.0014	(0.0129,1.836)	0.3466
	80	0.4974	0.4942	0.0011	(0.0131,1.811)	0.346
	100	0.4991	0.4982	0.0004	(0.0128,1.804)	0.3467
$\alpha = 3.5$	20	3.506	1.329	0.0057	(1.426,6.574)	3.332
	40	3.477	1.302	0.0054	(1.429,6.445)	3.31
	60	3.502	1.33	0.0034	(1.402,6.541)	3.326
	80	3.503	1.314	0.0032	(1.42,6.477)	3.316
	100	3.499	1.302	0.0029	(1.405,6.541)	3.328
$\beta = 3.5$	20	3.498	1.301	0.0059	(1.432,6.529)	3.347
	40	3.482	1.329	0.0059	(1.396,6.477)	3.317
	60	3.506	1.331	0.0036	(1.404,6.595)	3.347
	80	3.502	1.328	0.0035	(1.407,6.536)	3.333
	100	3.503	1.328	0.0032	(1.411,6.578)	3.342
$\lambda = 3.5$	20	3.498	1.326	0.0063	(1.403,6.532)	3.323
	40	3.511	1.328	0.0059	(1.404,6.582)	3.348
	60	3.501	1.322	0.0038	(1.411,6.514)	3.337
	80	3.502	1.324	0.0032	(1.408,6.555)	3.334
	100	3.499	1.318	0.0029	(1.41,6.512)	3.337

6 Real Data Application

This section presents a real data set analysis using the PNH distribution and further compares it with competing models, like the exponentiated-NH (ENH) (Lemonte, 2013), exponentiated Weibull (EW) (Mudholkar and

Table 4 Bladder Cancer Data Set

0.08	2.09	3.48	4.87	6.94	8.66	13.11	23.63	0.20	2.23	3.52	4.98	6.97	9.02	13.29
0.40	2.26	3.57	5.06	7.09	9.22	13.80	25.74	0.50	2.46	3.64	5.09	7.26	9.47	14.24
25.82	0.51	2.54	3.70	5.17	7.28	9.74	14.76	26.31	0.81	2.62	3.82	5.32	7.32	10.06
14.77	32.15	2.64	3.88	5.32	7.39	10.34	14.83	34.26	0.90	2.69	4.18	5.34	7.59	10.66
15.96	36.66	1.05	2.69	4.23	5.41	7.62	10.75	16.62	43.01	1.19	2.75	4.26	5.41	7.63
17.12	46.12	1.26	2.83	4.33	5.49	7.66	11.25	17.14	79.05	1.35	2.87	5.62	7.87	11.64
17.36	1.40	3.02	4.34	5.71	7.93	11.79	18.10	1.46	4.40	5.85	8.26	11.98	19.13	1.76
3.25	4.50	6.25	8.37	12.02	2.02	3.31	4.51	6.54	8.53	12.03	20.28	2.02	3.36	6.76
12.07	21.73	2.07	3.36	6.93	8.65	12.63	22.69							

Table 5 Monte Carlo Markov Chain results for the Bayesian analysis of the data set

Parameter↓	Estimate	SD	MC error	95% CI	Median
α	0.7721	0.7783	0.0036	(0.0189,2.863)	0.5339
β	1.666	1.672	0.0078	(0.0420,6.182)	1.152
λ	0.2021	0.2017	0.0009	(0.0052,0.7446)	0.1409

Srivastava, 1993), Marshall-Olkin Weibull (MOW) (Ghitany et al., 2005), BE, NH, EE and Weibull models. We estimate the model parameters by using the maximum likelihood method and compared the goodness-of-fit of the models using the Cramér–von Mises (W^*) and Anderson-Darling (A^*) statistics, which are described in detail by Chen and Balakrishnan (1995). In addition, we consider the Kolmogrov-Smirnov (K-S) statistic. In general, the smaller the values of these statistics, the better the fit to the data. The cdfs of the ENH, EW, MOW, BE and EE models are given by

$$
\begin{aligned}
&\text{ENH: } F_{ENH}(x;\beta,\alpha,\lambda) = \left(1 - \mathrm{e}^{1-(1+\lambda\, x)^{\alpha}}\right)^{\beta}, \quad x,\beta,\alpha,\lambda > 0,\\
&\text{EW: } F_{EW}(x;c,\alpha,\lambda) = \left(1 - \mathrm{e}^{-(x/\lambda)^{c}}\right)^{\alpha}, \quad x,c,\alpha,\lambda > 0,\\
&\text{MOW: } F_{MOW}(x;p,\beta,\lambda) = \left(1 - \mathrm{e}^{-(x/\lambda)^{\beta}}\right)\left[1 - (1-p)\,\mathrm{e}^{-(x/\lambda)^{\beta}}\right]^{-1},\\
&\qquad x,p,\beta,\lambda > 0,\\
&\text{BE: } F_{BE}(x;a,b,\lambda) = I_{1-\mathrm{e}^{-\lambda\, x}}(a,b), \quad x,a,b,\lambda > 0,\\
&\text{EE: } F_{EE}(x;\alpha,\lambda) = \left(1 - \mathrm{e}^{-\lambda\, x}\right)^{\alpha}, \quad x,\alpha,\lambda > 0,
\end{aligned}
$$

where $I_z(a,b)$ is the incomplete beta function ratio.

The data set has been taken from Lee and Wang (2013) and reproduced in Table 4, which represents the remission times (in months) of a random sample of 128 bladder cancer patients.

The box-plot of these observations is displayed in Figure 4(a), which indicates that the distribution is right-skewed. The TTT plot (Aarset, 1987) of these data is shown in Figure 4(b) and it is clear that it is first concave and then convex, which suggests an upside-down bathtub shaped failure rate. Accordingly, the PNH distribution could, in principle, be appropriate for modeling the current data set. The MLEs (with SEs in parentheses), A^*, W^* and K-S statistics are included in Table 6. All three goodness-of-fit statistics indicate that the PNH model provides the best fit. Further, the empirical and estimated survival curves and PP plot are shown in Figures 5(a) and 5(b) and also support this conclusion.

Table 6 MLEs, their standard errors (in parentheses) and goodness-of-fit statistics for the bladder cancer data

Distribution	Estimates			A^*	W^*	K-S
PNH	$\beta = 1.6362(0.2998)$	$\alpha = 0.7240(0.1560)$	$\lambda = 0.2011(0.1012)$	0.2422	0.0358	0.0405
ENH	$\beta = 1.6884(0.3646)$	$\alpha = 0.6371(0.1172)$	$\lambda = 0.3444(0.1752)$	0.2779	0.0421	0.0442
EW	$c = 0.6545(0.1346)$	$\alpha = 2.7942(1.2626)$	$\lambda = 3.3483(1.8911)$	0.2885	0.0436	0.0450
MOW	$p = 13.5390(1.7811)$	$\beta = 0.5445(0.1491)$	$\lambda = 0.9449(0.8414)$	0.8311	0.1417	0.0791
BE	$a = 1.1879(0.1352)$	$b = 4.0609(13.2406)$	$\lambda = 0.0306(0.0982)$	0.7154	0.1192	0.0738
NH	$\alpha = 0.9226(0.1514)$	$\lambda = 0.1216(0.0343)$		0.6741	0.11008	0.0919
EE	$\alpha = 1.2179(0.1488)$	$\lambda = 0.1211(0.0135)$		0.6033	0.1122	0.0725
W	$c = 1.0477(0.0675)$	$\alpha = 9.5585(0.8526)$		0.7863	0.1313	0.0699

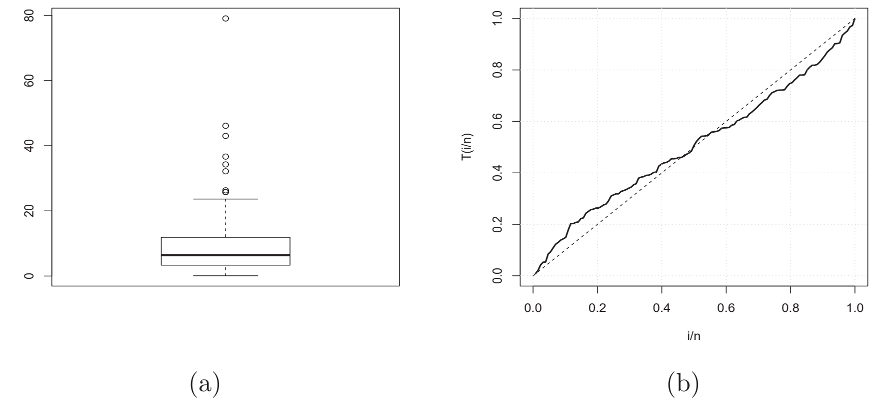


Figure 4 (a) Boxplot (b) TTT plot for the bladder cancer data.

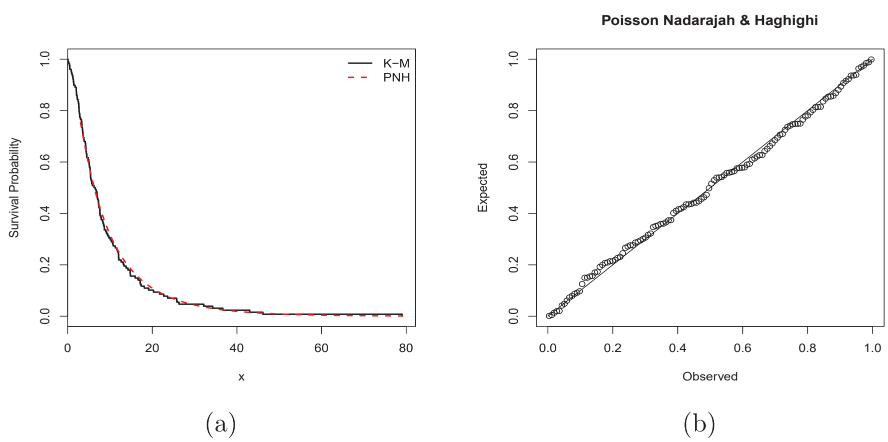


Figure 5 Bladder cancer data (a) empirical survival and estimated PNH survival function; (b) P-P plot.

7 Conclusion

In this article, we studied some basic statistical properties of the Poisson Nadarajah-Haghighi (PNH) distribution and estimated its parameters by eleven different methods of estimation, namely the maximum likelihood estimators, least squares and weighted least squares estimators, the maximum product of spacings estimators, the minimum spacing absolute distance estimators, the minimum spacing absolute-log distance estimators, Cramér-von-Mises estimators, Anderson-Darling and right-tail Anderson-Darling estimators and the Bayes estimators. Results of the simulation study showed that among frequentist estimators, WLS and MPS perform better than the other methods. However, the Bayesian is the best method. An application to a real data set is also presented as an illustration of the potentiality of the new model as compared to other existing models. It is expected the utility of the model in different fields, especially in survival analysis when hazard rate is decreasing, increasing, bathtub and upside-down bathtub shape. Further, it is also noticed that the performance of the MLEs is quite satisfactory. The use of the MLEs or Bayes estimators is recommended for practical purposes. In the future, record values can be analyzed assuming the PNH distribution.

Acknowledgments

The authors express their sincere thanks to the three reviewers and the editors for making some useful suggestions on an earlier version of this manuscript which resulted in this improved version.

References

Aarset, M. V. (1987). How to identify a bathtub hazard rate. *IEEE Transactions on Reliability*, R-36(1):106–108.

Ali, S., Dey, S., Tahir, M. H., and Mansoor, M. (2020a). A comparison of different methods of estimation for the flexible Weibull distribution. *Communications Faculty of Sciences University of Ankara Series A1 Mathematics and Statistics*, 69:794–814.

Ali, S., Dey, S., Tahir, M. H., and Mansoor, M. (2020b). Two-parameter logistic-exponential distribution: Some new properties and estimation methods. *American Journal of Mathematical and Management Sciences*, 39(3):270–298.

Ali, S., Dey, S., Tahir, M. H., and Mansoor, M. (2020c). Two-Parameter Logistic-Exponential Distribution: Some New Properties and Estimation Methods. *American Journal of Mathematical and Management Sciences*, 39(3):270–298.

Alizadeh, M., Afify, A. Z., Eliwa, M. S., and Ali, S. (2020). The odd log-logistic Lindley-G family of distributions: properties, Bayesian and non-Bayesian estimation with applications. *Computational Statistics*, 35:281–308.

Anderson, T. W. and Darling, D. A. (1952), Asymptotic Theory of Certain “Goodness of Fit” Criteria Based on Stochastic Processes, *The Annals of Mathematical Statistics*, 23(2):193–212.

https://doi.org/10.1214/aoms/1177729437

Chen, G. and Balakrishnan, N. (1995). A general purpose approximate goodness-of-fit test. *Journal of Quality Technology*, 27(2):154–161.

Cheng, R. C. H. and Amin, N. A. K. (1979). Maximum product of spacings estimation with application to the lognormal distribution, math.

Cheng, R. C. H. and Amin, N. A. K. (1983). Estimating parameters in continuous univariate distributions with a shifted origin. *Journal of the Royal Statistical Society. Series B (Methodological)*, 45(3):394–403.

Devroye, L. (1986). *Non-Uniform Random Variate Generation*. Springer-Verlag.

Dey, S., Al-Zahrani, B., and Basloom, S. (2017a). Dagum distribution: Properties and different methods of estimation. *International Journal of Statistics and Probability*, 6(2):74–92.

Dey, S., Ali, S., and Park, C. (2015). Weighted exponential distribution: properties and different methods of estimation. *Journal of Statistical Computation and Simulation*, 85(18):3641–3661.

Dey, S., Dey, T., Ali, S., and Mulekar, M. S. (2016). Two-parameter Maxwell distribution: Properties and different methods of estimation. *Journal of Statistical Theory and Practice*, 10(2):291–310.

Dey, S., Dey, T., and Kundu, D. (2014). Two-parameter Rayleigh distribution: Different methods of estimation. *American Journal of Mathematical and Management Sciences*, 33(1):55–74.

Dey, S., Kumar, D., Ramos, P. L., and Louzada, F. (2017b). Exponentiated Chen distribution: Properties and estimation. *Communications in Statistics – Simulation and Computation*, 46(10):8118–8139.

Dey, S., Raheem, E., and Mukherjee, S. (2017c). Statistical Properties and Different Methods of Estimation of Transmuted Rayleigh Distribution. *Revista Colombiana de EstadÃstica*, 40:165–203.

Dey, S., Raheem, E., Mukherjee, S., and Ng, H. K. T. (2017d). Two parameter exponentiated Gumbel distribution: properties and estimation with flood data example. *Journal of Statistics and Management Systems*, 20(2):197–233.

Eliwa, M. S., El-Morshedy, M., and Ali, S. (2020). Exponentiated odd Chen-G family of distributions: statistical properties, Bayesian and non-Bayesian estimation with applications. *Journal of Applied Statistics*, 0(0):1–27.

Ghitany, M. E., Al-Hussaini, E. K., and Al-Jarallah, R. A. (2005). Marshall–Olkin extended Weibull distribution and its application to censored data. *Journal of Applied Statistics*, 32(10):1025–1034.

Gupta, R. D. and Kundu, D. (1999). Theory & methods: Generalized exponential distributions. *Australian & New Zealand Journal of Statistics*, 41(2):173–188.

Kao, J. H. K. (1958). Computer methods for estimating Weibull parameters in reliability studies. *IRE Transactions on Reliability and Quality Control*, PGRQC-13:15–22.

Kao, J. H. K. (1959). A graphical estimation of mixed Weibull parameters in life-testing of electron tubes. *Technometrics*, 1(4):389–407.

Kenney, J. and Keeping, E. (1962). *Mathematics of statistics*. Number v. 2 in Mathematics of Statistics. Princeton: Van Nostrand.

Kundu, D. and Raqab, M. Z. (2005). Generalized Rayleigh distribution: different methods of estimations. *Computational Statistics & Data Analysis*, 49(1):187–200.

Lee, E. T. and Wang, J. W. (2013). *Statistical Methods for Survival Data Analysis*. Wiley Publishing, 4th edition.

Lehmann, E. and Casella, G. (2003). *Theory of Point Estimation*. Springer New York.

Lemonte, A. J. (2013). A new exponential-type distribution with constant, decreasing, increasing, upside-down bathtub and bathtub-shaped failure rate function. *Computational Statistics & Data Analysis*, 62:149–170.

MacDonald, P. D. M. (1971). Comment on "an estimation procedure for mixtures of distributions" by Choi and Bulgren. *Journal of the Royal Statistical Society. Series B (Methodological)*, 33(2):326–329.

Mansoor, M., Tahir, M. H., Alzaatreh, A., and Cordeiro, G. M. (2020a). The Poisson Nadarajah–Haghighi distribution: Properties and applications to lifetime data. *International Journal of Reliability, Quality and Safety Engineering*, 27(01):2050005.

Mansoor, M., Tahir, M. H., Cordeiro, G. M., Ali, S., and Alzaatreh, A. (2020b). The Lindley negative-binomial distribution: Properties, estimation and applications to lifetime data. *Mathematica Slovaca*, 70(4):917–934.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953). Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1092.

Moors, J. J. A. (1988). A quantile alternative for kurtosis. *Journal of the Royal Statistical Society. Series D (The Statistician)*, 37(1):25–32.

Mudholkar, G. S. and Srivastava, D. K. (1993). Exponentiated Weibull family for analyzing bathtub failure-rate data. *IEEE Transactions on Reliability*, 42(2):299–302.

Nadarajah, S. and Haghighi, F. (2011). An extension of the exponential distribution. *Statistics*, 45(6):543–558.

Nadarajaha, S. and Kotz, S. (2006). The beta exponential distribution. *Reliability Engineering & System Safety*, 91(6):689–697.

Ranneby, B. (1984). The maximum spacing method. an estimation method related to the maximum likelihood method. *Scandinavian Journal of Statistics*, 11(2):93–112.

Shafqat, M., Ali, S., Shah, I., and Dey, S. (2020). Univariate discrete Nadarajah and Haghighi distribution: Properties and different methods of estimation. *Statistica*, 80(3):301–330.

Swain, J. J., Venkatraman, S., and Wilson, J. R. (1988). Least-squares estimation of distribution functions in Johnson's translation system. *Journal of Statistical Computation and Simulation*, 29(4):271–297.

Tahir, M. H., Cordeiro, G. M., Ali, S., Dey, S., and Manzoor, A. (2018). The inverted Nadarajah–Haghighi distribution: estimation methods and applications. *Journal of Statistical Computation and Simulation*, 88(14):2775–2798.

Teimouri, M., Hoseini, S. M., and Nadarajah, S. (2013). Comparison of estimation methods for the Weibull distribution. *Statistics*, 47(1):93–109.

Torabi, H. (2008). A general method for estimating and hypotheses testing using spacings. *Journal of Statistical Theory and Practice*, 8(2):163–168.

Biographies

Sajid Ali is currently Assistant Professor at the Department of Statistics, Quaid-i-Azam University (QAU), Islamabad, Pakistan. He graduated (PhD Statistics) from Bocconi University, Milan, Italy. His research interest is focused on Bayesian inference, construction of new flexible probability distributions, time series analysis, and process monitoring.

Sanku Dey, M.Sc., Ph.D.: An Associate Professor in the Department of Statistics, St. Anthony's College, Shillong, Meghalaya, India. He has to his credit more than 220 research papers in journals of repute. He is a reviewer and associate editors of reputed international journals. He has a good number of contributions in almost all fields of Statistics viz., distribution theory, discretization of continuous distribution, reliability theory, multi-component stress-strength reliability, survival analysis, Bayesian inference, record statistics, statistical quality control, order statistics, lifetime performance index based on classical and Bayesian approach as well as different types of censoring schemes etc.

M. H. Tahir is currently Professor of Statistics, and Chair Department of Statistics at The University of Bahawalpur (IUB), Bahawalpur, Pakistan. He received BSc, MSc and PhD degree from IUB in 1988, 1990 and 2010, respectively. Dr Tahir has over 27 years of teaching experience to post-graduate classes, and has supervised 65 MPhil and 8 PhD students successfully. He has published more than 90 research papers in national and international journals, including Journal of Statistical Planning and Inference, Communications in Statistics-Theory and Methods, Communications in Statistics-Simulation and Computation, Journal of Statistical Computation and Simulation, Journal of Statistical Distributions and Applications, Journal of Statistical Theory and Applications. He is reviewer of more than 55 national and international statistical journals. Dr. Tahir's research interests include distribution theory, generalized classes of distributions, survival and lifetime data analysis, methods of estimation, and construction of experimental designs.

Muhammad Mansoor is Assistant Professor of Statistics at the Department of Statistics, Government Sadiq Egerton Graduate College, Bahawalpur, Pakistan. His current research focuses on generalizing statistical distributions arising from the hazard function. Other research areas include statistical inference of probability models, computational statistics, and regression analysis.