Thailand Statistician
January 2021; 19(1): 14-41
http://statassoc.or.th
Contributed paper

A Comprehensive Simulation Study to Compare Various Estimators of the Model Parameters, Model Mean, as well as Model Percentiles of a Two-Parameter Generalized Half-Normal Distribution (2P-GHND) with Applications

Matinee Sudsawat [a][b], Suntaree Unhapipat [b][c] and Nabendu Pal* [a]
[a] Department of Mathematics, University of Louisiana at Lafayette, Lafayette, Louisiana, USA.
[b] Department of Mathematics, Faculty of Science, Mahidol University, Bangkok, Thailand.
[c] Center of Excellence in Mathematics, Commission on Higher Education, Bangkok, Thailand.
*Corresponding author; e-mail: nabendu.pal@louisiana.edu

Received: 18 August 2019
Revised: 13 October 2019
Accepted: 5 March 2020

Abstract

This work deals with studying various point estimators of the model parameters, the model mean, as well as the model percentiles of a two-parameter generalized half normal distribution (2P-GHND). First, we study three types of estimators of the model parameters, namely - the method of moments estimators (MMEs), the maximum likelihood estimators (MLEs), and the ordinary regression estimators (OREs). Then, these three methods are used to obtain the corresponding estimators of the model mean as well as the model percentiles. The estimators have been compared in terms of relative bias (RB) and relative mean squared error (RMSE). Though our primary objective here is to study the small sample behaviour of the estimators, we have also studied the asymptotic behaviour of the MLEs. It has been shown that the MLEs perform far better than the other types of estimators for sample sizes up to 25. For larger sample sizes, all the estimators have nearly similar behaviour. Also, for the MLEs of all the parameters considered in this study, their MSEs can be approximated fairly well by the respective asymptotic variances obtained from the Fisher information matrix. Finally, we provide asymptotic interval estimates of all the parameters considered here, and show the goodness of fit of 2P-GHND over other commonly used skewed distributions for two real-life datasets.

Keywords: Shape parameter, scale parameter, relative bias, relative mean squared error, asymptotic variance.

1. Introduction

The Gaussian or the normal distribution is the most widely used probability distribution to model real-life datasets. However, if a dataset shows a positively skewed pattern, and/or bounded below by zero, then the normality assumption may not be appropriate for such data. That is why in many biological and engineering applications we see usage of positively skewed distributions, such as - lognormal, Weibull, gamma, etc. A competitor to the aforementioned positively skewed distributions is the two parameter generalized half normal distribution (or, 2P-GHND) which arises as a generalization of a folded form of a normal distribution centered at zero (i.e., half normal distribution).

Though the above-mentioned positively skewed distributions have been studied extensively in the literature, and have been used widely to model real-life datasets, relatively less has been done for the 2P-GHND, and this is the motivation behind this work.

In the following we briefly describe the 2P-GHND to be considered in this study and its basic properties.

1.1. Two-parameter GHND (2P-GHND) and basic properties

A random variable X is said to follow a two-parameter GHND (or 2P-GHND) with shape and scale parameters δ and σ respectively, provided its pdf is given as

$$f(x|\delta,\sigma) = \sqrt{2/\pi}(\delta/x)(x/\sigma)^{\delta} exp(-(x/\sigma)^{2\delta}/2), \tag{1}$$

where $\delta > 0$, $\sigma > 0$ and $x > 0$. The cdf of 2P-GHND is given as

$$P(X \leq x) = F(x) = \int_0^x f(x|\delta,\sigma)dx = 2\Phi[(x/\sigma)^{\delta}] - 1, \tag{2}$$

where Φ represents the standard normal cdf. The 2P-GHND model is appropriate for nonnegative observations, such as the life time of equipments, amount of rainfall per unit of time, etc.

For any integer $k \geq 1$, the k^{th} raw moment of 2P-GHND has the following expression

$$E(X^k) = \sigma^k \sqrt{(2^{k/\delta}/\pi)}\Gamma((k+\delta)/(2\delta)), \tag{3}$$

where $\Gamma(c)$ is the regular gamma function evaluated at c. Thus, the first two moments of 2P-GHND are

$$E(X) = \sigma\sqrt{(2^{1/\delta}/\pi)}\Gamma((1+\delta)/(2\delta)), \quad E(X^2) = \sigma^2\sqrt{(2^{2/\delta}/\pi)}\Gamma((2+\delta)/(2\delta)). \tag{4}$$

The objective of this work is to undertake a comprehensive simulation study of various estimators of the model parameters, i.e., δ (the shape parameter) and σ (the scale parameter), as well as two other important parameters, namely - η (the model mean), and ξ_p (the p^{th} percentile of the model, $0 < p < 1$).

The expression (1) clearly shows how the model parameters δ and σ are involved in the pdf. Further, from (4) it is seen that

$$\eta = E(X) = \sigma\sqrt{(2^{1/\delta}/\pi)}\Gamma((1+\delta)/(2\delta)),$$

and from (2) it is seen that the p^{th} percentile of the model is given by

$$\xi_p = \sigma\{\Phi^{-1}((p+1)/2)\}^{1/\delta}.$$

In the following we provide a brief literature survey on 2P-GHND.

1.2. Some historical background

The properties of half-normal distribution and its generalizations have been studied by a few researchers in the recent past. Pewsey (2002) gave a nice historical development of a variant of 2P-GHND with location and scale parameters (which is different from (1) considered here), and studied the point as well as interval estimators of the parameters for large samples. Pewsey (2004) extented this study further by proposing bias corrected point estimators of the model parameters which tend to perform better than the original MLEs both in terms of bias and MSE. Cooray and Ananda (2008) considered the 2P-GHND (with pdf (1)), studied various distributional as well as characterization properties of the distribution, and proposed interval estimates which were than studied for small to moderate sample sizes for a very limited combination of the model parameters. They fitted the 2P-GHND model to static fatigue life of objects like glass, ceramics, etc. under a constant stress, and compared the model fitting with that under a few other models, such as - gamma, lognormal, Weibull and Birnbaum-Saunders distributions by using four criteria, which are the value of the maximum log likelihood function, Kolmogorov-Smirnov statistic, Anderson-Darling statistic, and chi-square goodness of fit statistic. Gómez and Vidal (2016) considered a completely different type of generalization of the half-normal distribution which is beyond the scope of our study.

Note that the existing works do provide some simulation results which are limited in terms of scope and breadth, and it is not easy to decipher any meaningful trend. Therefore, in this work we have undertaken a comprehensive study of various estimators not only of the model parameters, but also of the the model mean as well as the model percentiles which were omitted in the previous works.

It is worth pointing out that Stacy (1962) introduced a very general class of skewed distributions, called generalized gamma distribution (GGD), for a nonnegative random variable, with the following pdf

$$f(x|a,d,p) = (p/a^d)x^{d-1}exp(-(x/a)^p)/\Gamma(d/p), \tag{5}$$

where $a > 0, d > 0, p > 0$, and $x > 0$. The above GGD pdf includes many well-known and widely used pdfs as special cases. For example, $p = 1$ reduces (5) into the regular gamma pdf (with scale parameter a and shape parameter d). Similarly, with $p = d$ makes (5) boil down to the regular Weibull pdf (with scale parameter a, and shape parameter d). As a result, the above GGD model includes exponential, chi-square, Rayleigh distributions as special cases. Further, by taking $p = 2\delta$, $d = \delta$ and $a = 2^{1/(2\delta)}\sigma$, one can arrive at the 2P-GHND from (5). Thus 2P-GHND is also a special case of GGD. However, in spite of being introduced more than five decades ago, GGD did not gain much popularity due to challenges with estimation of parameters, and subsequent sampling distributions. Any attempt to find the MLE of the three model parameters of GGD gets stymied by the system of three highly nonlinear equations (which may result into multiple solutions, or no solution at all). The computational challenges with GGD have also been acknowledged by Cooray and Ananda (2008) (see their Subsection 3.1). Hence, Stacy and Mihram (1965) suggested estimating the parameters using the method of moments. But to best of our knowledge, no comprehensive study exists comparing the goodness of the aforementioned estimators, and/or modifying the MLE approach. As a result, research has proliferated on specific subfamilies of GGD, such as Weibull and gamma, and how to make use of these specific subfamilies in real-life applications. In the same vein we study 2P-GHND in this paper which has received less attention than Weibull and Gamma but yet found to be useful for some datasets as shown later.

Even though 2P-GHND is a member of GGD, unlike other members like regular gamma it is not always positively skewed (a property shared by Weibull though). This interesting feature had eluded the attention of earlier researches. Note that the measure of skewness (MOS) of 2P-GHND, denoted by γ_0, is given by

$$\gamma_0 = E(X-\eta)^3/(E(X-\eta)^2)^{(3/2)} = w_1/w_2,$$

where $w_1 = \sqrt{(2^{3/\delta}/\pi)}\Gamma\left((3+\delta)/(2\delta)\right) + 2\left(\sqrt{(2^{1/\delta}/\pi)}\Gamma\left((1+\delta)/(2\delta)\right)\right)^3$
$-3\sqrt{(2^{3/\delta}/\pi^2)}\Gamma\left((1+\delta)/(2\delta)\right)\Gamma\left((2+\delta)/(2\delta)\right)$,
and $w_2 = \left(\sqrt{(2^{2/\delta}/\pi)}\Gamma\left((2+\delta)/(2\delta)\right) - \left(\sqrt{(2^{1/\delta}/\pi)}\Gamma\left((1+\delta)/(2\delta)\right)\right)^2\right)^{(3/2)}$.

Thus, γ_0 is a function of δ only, and its behavior is shown in the following Figure 1. Interestingly γ_0 is positive only when δ is < 2.175, and negative when δ is > 2.175. In order words, 2P-GHND goes from positively skewed to negatively skewed as δ increases from near 0, and crosses the value 2.175. This gives 2P-GHND some flexibility over other strictly positively skewed distributions in fitting real-life data.

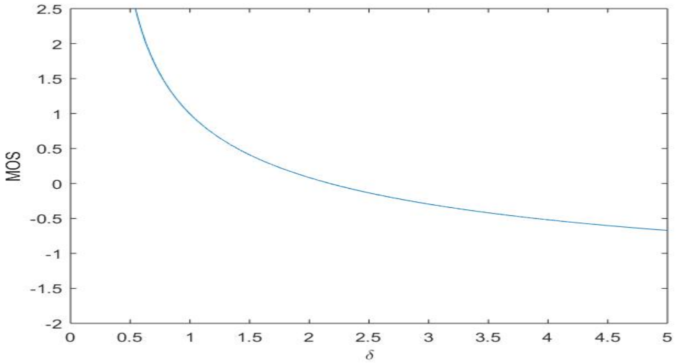


Figure 1 Plot of MOS against the shape parameter

The rest of the paper is organized as follows. Section 2 deals with estimation of the model parameters δ and σ. Here we present three types of estimators and explain some of their properties. In Section 3 we consider estimation of the model mean η, and Section 4 presents the estimation of the model percentile ξ_p. An important component of this study is to see for what sample size the asymptotic behaviour of the above mentioned estimators take hold. Also, the interval estimation of the relevant parameters are shown in Section 5. In Section 6 we present two datasets which can be modelled by 2P-GHND satisfactorily over a few other two parameter positively skewed distributions. For convenience all plots and graphs have been relegated to the Appendix (after the list of reference). Finally, the paper ends with some practical observations and recommendations.

2. Estimation of the Model Parameters (δ and σ)

Suppose $X_1, X_2, ..., X_n$ are i.i.d. following 2P-GHND with pdf (1). Though the expressions of the estimators in Subsections 2.1 and 2.2 are not new, still we have presented them here for completeness.

2.1. Method of moments estimators (MME)

We define the k^{th} sample raw moment m_k as

$$m_k = \sum_{i=1}^{n} X_i^k / n, \;\; k \geq 1.$$

The method of moments estimators (MMEs) of δ and σ are found by equating the above first two sample raw moments with their population counterparts from (3), i.e.,

$$m_1 = \sigma \sqrt{(2^{1/\delta}/\pi)} \Gamma((1+\delta)/(2\delta)), \tag{6}$$

$$m_2 = \sigma^2 \sqrt{(2^{2/\delta}/\pi)} \Gamma((2+\delta)/(2\delta)). \tag{7}$$

From equations (6) and (7), the MME of δ, i.e., $\widehat{\delta}_{MME}$, can be found by solving the following equation numerically

$$(m_1^2/m_2) = (\Gamma((1+\widehat{\delta}_{MME})/(2\widehat{\delta}_{MME})))^2/(\pi\Gamma((2+\widehat{\delta}_{MME})/(2\widehat{\delta}_{MME}))),$$

and then replace δ by $\widehat{\delta}_{MME}$ in equation (6) to get the MME of σ, i.e., get $\widehat{\sigma}_{MME}$, by the expression

$$\widehat{\sigma}_{MME} = m_1/[\sqrt{(2^{1/\widehat{\delta}_{MME}}/\pi)}\Gamma((1+\widehat{\delta}_{MME})/(2\widehat{\delta}_{MME}))].$$

Remark 1 It is easy to see that the probability distributions of $\widehat{\delta}_{MME}$ and $(\widehat{\sigma}_{MME}/\sigma)$ are free from σ, and depend only on δ (apart from n).

2.2. Maximum likelihood estimators (MLE)

The log-likelihood function (L_*) of 2P-GHND with n observations is

$$L_*(\delta, \sigma|\mathbf{X}) = (n/2)\ln(2/\pi) + n\ln(\delta) - n\delta\ln(\sigma) + (\delta - 1)\sum_{i=1}^{n}(\ln(X_i)) - (\sum_{i=1}^{n}(X_i/\sigma)^{2\delta}/2).$$

Differentiating $L_*(\delta, \sigma|\mathbf{X})$ with respect to δ and σ, and then equating them with 0 gives the MLEs of δ and σ as follows. First obtain $\widehat{\delta}_{MLE}$ by solving

$$1/\widehat{\delta}_{MLE} = (\sum_{i=1}^{n} X_i^{2\widehat{\delta}_{MLE}} \ln X_i)/(\sum_{i=1}^{n} X_i^{2\widehat{\delta}_{MLE}}) - \sum_{i=1}^{n}(\ln X_i)/n,$$

and then obtain the MLE of σ, i.e., $\widehat{\sigma}_{MLE}$ as

$$\widehat{\sigma}_{MLE} = [\sum_{i=1}^{n} X_i^{2\widehat{\delta}_{MLE}}/n]^{1/(2\widehat{\delta}_{MLE})}.$$

Remark 2 It is easy to see that the probability distributions of $\widehat{\delta}_{MLE}$ and $(\widehat{\sigma}_{MLE}/\sigma)$ are also free from σ, and depend only on δ (apart from n).

2.3. Ordinary regression estimators (ORE)

The third method is the ordinary regression estimation through the least squares method. Writing $X_{(1)} \leq X_{(2)} \leq ... \leq X_{(n)}$ as the ordered observations, compute $c_{i/n}$ defined as

$$c_{i/n} = \Phi^{-1}((1 + i/n)/2),$$

where $i = 1, 2, ..., n$, and Φ^{-1} is the inverse of Φ as defined in (2).

Since $X_{(i)}$ is the empirical $(i/n)100^{\text{th}}$ percentile value, it is matched with the corresponding population percentile value (also see Section 4 later), i.e.,

$$X_{(i)} = \sigma({c_{i/n}}^{1/\delta}),$$

i.e.,

$$\ln X_{(i)} = \ln(\sigma) + (1/\delta)\ln(c_{i/n}).$$

Substituting $\sigma_0 = \ln(\sigma)$ and $(1/\delta) = \delta_0$, we get

$$\ln X_{(i)} = \sigma_0 + \delta_0 \ln(c_{i/n}).$$

Now, by the simple regression method, get the estimated parameters $\widehat{\delta}_0$ and $\widehat{\sigma}_0$ as follows

$$\widehat{\delta}_0 = [\sum_{i=1}^{n}(u_i - \overline{u}_i)(v_i - \overline{v}_i)]/[\sum_{i=1}^{n}(v_i - \overline{v}_i)^2] \qquad \text{and} \qquad \widehat{\sigma}_0 = \overline{u}_i - \widehat{\delta}_0\overline{v}_i,$$

where $u_i = \ln X_{(i)}$ and $v_i = \ln(c_{i/n})$. Finally, the estimated parameters, based on the ordinary regression estimation (ORE), are found as

$$\widehat{\delta}_{ORE} = 1/\widehat{\delta}_0 \qquad \text{and} \qquad \widehat{\sigma}_{ORE} = exp(\widehat{\sigma}_0).$$

2.4. Simulation to compare estimators of δ and σ

In this simulation study, the estimators are compared by using the relative bias (RB) and the relative mean squared error (RMSE). The RB and the RMSE of any generic estimator, say $\widehat{\theta}$ of a parameter θ, is defined as

$$RB(\widehat{\theta}) = Bias(\widehat{\theta})/\theta \qquad \text{and} \qquad RMSE(\widehat{\theta}) = MSE(\widehat{\theta})/\theta^2.$$

Even though it is customary to use the exact bias and exact MSE of an estimator, say $\widehat{\theta}$, of a parameter θ, where $Bias(\widehat{\theta}) = E(\widehat{\theta} - \theta)$ and $MSE(\widehat{\theta}) = E(\widehat{\theta} - \theta)^2$, the RB and RMSE are more informative about the true performance of $\widehat{\theta}$ at θ. For example, a bias of 0.1 may not be bad for estimating θ at $\theta = 10$, but it can be bad if $\theta = 1.0$. This difference gets reflected in RB and RMSE.

Note that Bias (or MSE) of an estimator is an expectation of an expression, and hence it has been approximated by the average of a large number (say, M) of replicated values of the same expectant for a fixed sample size with an input value of the parameters. W.l.g., we have used $\sigma = 1$ throughout this study, and M has been taken as 10^5.

Remark 3 The Figure 2 shows the RB of three δ estimators, i.e., $\widehat{\delta}_{MME}$, $\widehat{\delta}_{MLE}$ and $\widehat{\delta}_{ORE}$. Note that RB of $\widehat{\delta}_{MLE}$ and $\widehat{\delta}_{ORE}$ are almost constant with respect to δ for all n. Although, the RB of $\widehat{\delta}_{MME}$ is higher than those of other two for $\delta < 1$, it decreases monotonically, and then becomes almost flat when δ is greater than 1.0. Interestingly, among these three estimators, $\widehat{\delta}_{ORE}$ has the smallest RB. All the three RBs approach 0, when n increases. Therefore, the estimator having the least bias is $\widehat{\delta}_{ORE}$. On the other hand, Figure 3 shows the RMSE of the three δ estimators, and it is observed that the best overall performance is shown by $\widehat{\delta}_{MLE}$, closely followed by $\widehat{\delta}_{ORE}$. Only for $n = 5$, $\widehat{\delta}_{MME}$ shows some good performance in terms of RMSE when $\delta > 0.5$. Taking all these aspects into consideration, we therefore conclude that $\widehat{\delta}_{MLE}$ is the most suitable estimator for δ.

Remark 4 For estimating σ, Figures 4 and 5 show the behaviour of $\widehat{\sigma}_{MME}$, $\widehat{\sigma}_{MLE}$ and $\widehat{\sigma}_{ORE}$ in terms of RB and RMSE respectively. Note that both $\widehat{\sigma}_{MME}$ and $\widehat{\sigma}_{MLE}$ have nonnegative RB, and this is almost zero for $\delta > 0.5$ for all n. On the other hand, $\widehat{\sigma}_{ORE}$ is mostly negatively biased, i.e., $\widehat{\sigma}_{ORE}$ underestimates σ. All the RBs are approaching zero as n increases. In terms of RMSE, $\widehat{\sigma}_{MLE}$ is certainly better than the other two for $\delta > 0.5$. However, as n increases, $\widehat{\sigma}_{MLE}$ has the overall best performance.

Based on our simulation study, in terms of overall performance, the MLE stands out as the best method for estimating both δ and σ. In the following we further study the asymptotic performance of $\widehat{\delta}_{MLE}$ and $\widehat{\sigma}_{MLE}$.

2.5. Asymptotic behaviour of the MLE for δ and σ

Since 2P-GHND is a member of the exponential family of distributions, the standard large sample theory says that $(\widehat{\delta}_{MLE}, \widehat{\sigma}_{MLE})'$ has an asymptotic bivariate normal distribution. To be specific, as $n \to \infty$

$$\begin{pmatrix} \widehat{\delta}_{MLE} \\ \widehat{\sigma}_{MLE} \end{pmatrix} \longrightarrow^d N_2\left(\begin{pmatrix} \delta \\ \sigma \end{pmatrix}, I^{-1}_{2\times 2}\right),$$

where I^{-1} is the inverse of the Fisher information matrix of 2P-GHND (the notation $\longrightarrow^d$ implies convergence in distribution). The Fisher information matrix of 2P-GHND is given as Mazucheli et al. (2018)

$$I(\delta, \sigma) = n\begin{bmatrix} (1/(2\delta^2))[\pi^2/2 - 2 + (2 - \ln(2) - \gamma)^2] & -(2 - \ln(2) - \gamma)/\sigma \\ -(2 - \ln(2) - \gamma)/\sigma & 2(\delta/\sigma)^2 \end{bmatrix},$$

where $\gamma = 0.5772...$ is the Euler's constant. Then, the inverse of the Fisher information matrix of 2P-GHND is presented as follows

$$I^{-1}(\delta,\sigma|x) = (1/n)(2\sigma^2/(\pi^2-4))I^{2\times 2}, \tag{8}$$

where

$$I^{2\times 2} = \begin{bmatrix} 2(\delta/\sigma)^2 & (2-\ln(2)-\gamma)/\sigma \\ (2-\ln(2)-\gamma)/\sigma & (1/(2\delta^2))[\pi^2/2-2+(2-\ln(2)-\gamma)^2] \end{bmatrix}.$$

Therefore, the asymptotic variance (AV) of $\widehat{\delta}_{MLE}$ and $\widehat{\sigma}_{MLE}$ are given as follows.

$$AV(\widehat{\delta}_{MLE}) = [4/(n(\pi^2-4))](\delta^2), \text{ and} \tag{9}$$

$$AV(\widehat{\sigma}_{MLE}) = [\sigma^2/(n(\pi^2-4))][\pi^2/2-2+(2-\ln(2)-\gamma)^2](1/\delta^2).$$

The comparison between the actual MSE and the AV of $\widehat{\delta}_{MLE}$ has been done in terms of RMSE and relative AV (i.e., RAV), defined as

$$RAV(\sqrt{n}\widehat{\delta}_{MLE}) = nAV(\widehat{\delta}_{MLE})/\delta^2 = [4/(\pi^2-4)], \text{ and}$$

$$RMSE(\sqrt{n}\widehat{\delta}_{MLE}) = nMSE(\widehat{\delta}_{MLE})/\delta^2,$$

which gives a true comparative picture irrespective of the sample size and the true value of δ. The same comparison has been done for $\widehat{\sigma}_{MLE}$ also. For the asymptotic variance of $\widehat{\delta}_{MLE}$ and $\widehat{\sigma}_{MLE}$, Figure 6 shows how the RMSE of $(\sqrt{n}\widehat{\delta}_{MLE})$ is getting closer to RAV of $(\sqrt{n}\widehat{\delta}_{MLE})$ as n increases. Similarly, Figure 7 shows the comparison between RMSE$(\sqrt{n}\widehat{\sigma}_{MLE})$ and RAV$(\sqrt{n}\widehat{\sigma}_{MLE})$ as a function of δ for various values of n with $\sigma = 1$ (w.l.g.).

Next we consider the association between $\widehat{\delta}_{MLE}$ and $\widehat{\sigma}_{MLE}$ by using the asymptotic correlation coefficient (ACC) as follows

$$ACC(\widehat{\delta}_{MLE},\widehat{\sigma}_{MLE}) = (2-\ln(2)-\gamma)/(\pi^2/2-2+(2-\ln(2)-\gamma)^2)^{(1/2)} \approx 0.391849. \tag{10}$$

Therefore, ACC of $\widehat{\delta}_{MLE}$ and $\widehat{\sigma}_{MLE}$ is a constant for lange n. In the following subsection we try to make use of the asymptotic correlation between $\widehat{\delta}_{MLE}$ and $\widehat{\sigma}_{MLE}$ for estimating δ and σ.

2.6. On asymptotic correlation between $\widehat{\delta}_{MLE}$ and $\widehat{\sigma}_{MLE}$

From the expression (10), it is seen that there is a positive asymptotic linear association between $\widehat{\delta}_{MLE}$ and $\widehat{\sigma}_{MLE}$ which takes effect for very large n and it is free form δ. How this association between $\widehat{\delta}_{MLE}$ and $\widehat{\sigma}_{MLE}$ takes place for fixed sample sizes can be seen from the simulated scatter plots based on 10^5 replicated values of $(\widehat{\delta}_{MLE},\widehat{\sigma}_{MLE})$, as presented in Figures 8 (for $n = 50$) and 9 (for $n = 100$) for various values of δ.

The whole idea of the above scatter plots is to see whether the association looks linear or not, though it may be dependent on δ because in real - life problems we have n finite. The Figures 8 and 9 do indicate linear association which varies mildly with respect to δ.

Based on our observations from the scatterplots, one can suggest a linear relationship between $(\widehat{\delta}_{MLE},\widehat{\sigma}_{MLE})$ for large n as

$$\widehat{\sigma}_{MLE} \approx a + b\,\widehat{\delta}_{MLE}, \tag{11}$$

where $a = \sigma - b\delta$ and $b \approx (0.391849)(AV(\widehat{\sigma}_{MLE})/AV(\widehat{\delta}_{MLE}))^{(1/2)} = (0.364819)(\sigma/\delta^2)$.

From (11), then

$$\widehat{\sigma}_{MLE} \approx (\sigma - b\delta) + b\,\widehat{\delta}_{MLE},$$

i.e.,

$$\widehat{\sigma}_{MLE} - \sigma \approx b(\widehat{\delta}_{MLE} - \delta).$$

In term of Bias of $\widehat{\delta}_{MLE}$ and $\widehat{\sigma}_{MLE}$,

$$Bias(\widehat{\sigma}_{MLE}) \approx bBias(\widehat{\delta}_{MLE}) = (0.364818577)(\sigma/\delta^2)Bias(\widehat{\delta}_{MLE}).$$

The above asymptotic bias relationship between $\widehat{\delta}_{MLE}$ and $\widehat{\sigma}_{MLE}$ gives us an idea of deriving bias corrected estimators of δ and σ for finite, yet large n. In the following we present a new set of bias corrected estimators through the Jackknife method (Efron (1982)).

We need to get estimators of the above two biases which can be obtained through the Jackknife (JK) method. Let us remove the observation X_i from the sample, i.e., use $X_1, ..., X_{i-1}, X_{i+1}, ..., X_n$ to get the MLEs of δ and σ as $\widehat{\delta}^{(i)}_{MLE}$ and $\widehat{\sigma}^{(i)}_{MLE}$ for $i = 1, 2, ..., n$ based on $(n-1)$ observations. Then the δ and σ estimators of the JK method is

$$\widehat{\delta}^{(\cdot)} = \sum_{i=1}^{n} \widehat{\delta}^{(i)}_{MLE}/n \qquad \text{and} \qquad \widehat{\sigma}^{(\cdot)} = \sum_{i=1}^{n} \widehat{\sigma}^{(i)}_{MLE}/n,$$

then the JK biases based on $\widehat{\delta}^{(\cdot)}$ and $\widehat{\sigma}^{(\cdot)}$ are defined as

$$\widehat{Bias}_{JK}(\widehat{\delta}_{MLE}) = (n-1)(\widehat{\delta}^{(\cdot)} - \widehat{\delta}_{MLE}),$$

and

$$\widehat{Bias}_{JK}(\widehat{\sigma}_{MLE}) = (n-1)(\widehat{\sigma}^{(\cdot)} - \widehat{\sigma}_{MLE}).$$

Hence, a new estimator of σ can be proposed by replacing $Bias(\widehat{\delta}_{MLE})$ and $Bias(\widehat{\sigma}_{MLE})$ by $\widehat{Bias}_{JK}(\widehat{\delta}_{MLE})$ and $\widehat{Bias}_{JK}(\widehat{\sigma}_{MLE})$ respectively, which produces the new estimator as

$$\widehat{\sigma}_{new} = (\widehat{\delta}_{MLE})^2(\widehat{Bias}_{JK}(\widehat{\sigma}_{MLE})/\widehat{Bias}_{JK}(\widehat{\delta}_{MLE}))(1/(0.364818577)).$$

Similarly, a new estimator of δ is found as

$$\widehat{\delta}_{new} = [(0.364818577)(\widehat{\sigma}_{MLE})(\widehat{Bias}_{JK}(\widehat{\delta}_{MLE})/\widehat{Bias}_{JK}(\widehat{\sigma}_{MLE}))]^{(1/2)}.$$

Figures 10 and 12 present the RB of $\widehat{\delta}_{new}$ and $\widehat{\sigma}_{new}$ which are less than the RB of $\widehat{\delta}_{MLE}$ and $\widehat{\sigma}_{MLE}$, but do not get close to zero. It appears that the JK method overkills by reducing bias too much. However, the MSE of $\widehat{\delta}_{new}$ and $\widehat{\sigma}_{new}$ are still higher than MLE estimator shown in Figures 11 and 13.

In short, by using the jackknife method, we obtain the new estimators by correcting the bias, however these new estimator still have higher RMSE. Therefore, the MLE is still the best method for estimating both δ and σ.

3. Estimation of the model mean (η)

Let η be the model mean of 2P-GHND, that is defined by

$$\eta = E(X) = \sigma\sqrt{(2^{1/\delta}/\pi)}\Gamma((1+\delta)/(2\delta)) = g(\delta, \sigma).$$

In Subsections 2.4 - 2.5 we have noted that the MLE approach is the overall best method of estimation, so we are going to study the asymptotic behavior of $\widehat{\eta}_{MLE}$, apart from $\widehat{\eta}_{MME} = \overline{X}$. It is easy to see that

$$E(\widehat{\eta}_{MME}) = E(\overline{X}) = (1/n)(nE(X_i)) = \sigma\sqrt{(2^{1/\delta}/\pi)}\Gamma((1+\delta)/(2\delta)) = \eta.$$

Therefore, $\widehat{\eta}_{MME}$ is an unbias estimator of η, i.e., Bias$(\widehat{\eta}_{MME}) = E(\widehat{\eta}_{MME} - \eta) = 0$, and the variance of $\widehat{\eta}_{MME}$ is

$$V(\widehat{\eta}_{MME}) = V(\overline{X}) = (1/n^2)(nV(X_i)) = (1/n)(V(X_i)),$$

where $V(X_i) = E(X_i^2) - (E(X_i))^2$. From Equations (4), the variance of X_i (i.e., $V(X_i)$) is

$$\begin{aligned} V(X_i) &= \sigma^2\sqrt{(2^{2/\delta}/\pi)}\Gamma((2+\delta)/(2\delta)) - [\sigma\sqrt{(2^{1/\delta}/\pi)}\Gamma((1+\delta)/(2\delta))]^2 \\ &= \sigma^2\sqrt{(2^{2/\delta}/\pi)}[\Gamma((2+\delta)/(2\delta)) - (1/\sqrt{\pi})(\Gamma((1+\delta)/(2\delta)))^2]. \end{aligned}$$

Hence,

$$V(\widehat{\eta}_{MME}) = (1/n)(\sigma^2\sqrt{(2^{2/\delta}/\pi)})[\Gamma((2+\delta)/(2\delta)) - (1/\sqrt{\pi})(\Gamma((1+\delta)/(2\delta)))^2)].$$

Therefore, the MSE of $\widehat{\eta}_{MME}$ is

$$MSE(\widehat{\eta}_{MME}) = E((\widehat{\eta}_{MME} - \eta)^2) = V(\widehat{\eta}_{MME}),$$

and the RMSE of $\widehat{\eta}_{MME}$ is

$$RMSE(\sqrt{n}\widehat{\eta}_{MME}) = (nMSE(\widehat{\eta}_{MME}))/\eta^2,$$

i.e.,

$$RMSE(\sqrt{n}\widehat{\eta}_{MME}) = [(\sqrt{\pi}(\Gamma((2+\delta)/(2\delta))))/(\Gamma((1+\delta)/(2\delta)))^2] - 1.$$

Next, we consider the asymptotic behaviour of $\widehat{\eta}_{MLE}$ by using its asymptotic properties. As $n \to \infty$,

$$\left(\ \widehat{\eta}_{MLE} \ \right) \longrightarrow^d N\left(g(\delta,\sigma), \nabla g(\delta,\sigma)' I^{-1} \nabla g(\delta,\sigma)\right),$$

where I^{-1} is a inverse of the fisher information matrix of 2P-GHND (see (8)), and

$$\nabla g(\delta,\sigma) = [(\partial/\partial\delta)g(\delta,\sigma), (\partial/\partial\sigma)g(\delta,\sigma)]'.$$

Therefore, the asymptotic mean of $\widehat{\eta}_{MLE}$ is

$$E(\widehat{\eta}_{MLE}) = g(\delta,\sigma) = \sigma\sqrt{(2^{1/\delta}/\pi)}\Gamma((1+\delta)/(2\delta)) = \eta,$$

and

$$\begin{aligned} AV(\sqrt{n}\widehat{\eta}_{MLE}) = {} & (1/(\pi\delta^2))[(\sigma^2 2^{1/\delta})/(\pi^2-4)](\Gamma((1+\delta)/(2\delta)))^2 \\ & [(\psi((1+\delta)/(2\delta)) + 3\ln(2) - 4 + 2\gamma)((\psi((1+\delta)/(2\delta)) + \ln(2)) \\ & + (\pi^2/2 - 2 + (2 - \ln(2) - \gamma)^2)]. \end{aligned} \quad (11)$$

The asymptotic behaviour of the MLE for η is shown in Figure 14 along with $RMSE(\sqrt{n}\widehat{\eta}_{MME})$. The RMSEs of $\widehat{\eta}_{MLE}$ are close to RAV of $\widehat{\eta}_{MLE}$ when n and δ increase.
Note that $RAV(\sqrt{n}\widehat{\eta}_{MME}) = AV(\sqrt{n}\widehat{\eta}_{MME})/(\eta^2)$.

4. Estimation of the Model p^{th} Percentile (ξ_p)

For any $p \in (0,1)$, the $(100p)^{\text{th}}$ percentile is ξ_p which satisfies the following equation (from (2))

$$2\Phi[(\xi_p/\sigma)^\delta] - 1 = p,$$

i.e.,

$$\xi_p = \sigma(c_p)^{1/\delta},$$

where $c_p = \Phi^{-1}((p+1)/2)$. In order to estimate ξ_p we use the estimated model parameters from one of the estimation methods discussed earlier which gives the expression of ξ_p estimator as

$$\widehat{\xi}_p = \widehat{\sigma}(c_p)^{1/\widehat{\delta}},$$

where $(\widehat{\delta}, \widehat{\sigma})$ can come from either MME, or MLE, or ORE. Thus we get estimates of ξ_p as $\widehat{\xi}_{p(MME)}$, $\widehat{\xi}_{p(MLE)}$ and $\widehat{\xi}_{p(ORE)}$, respectively. We have studied the performance of the above three estimators of ξ_p for $p = \{0.01, 0.05, 0.10, 0.50, 0.95, 0.99\}$ and $n = 5, 10, 25$ and 50. However, in this paper, we present the representative results for $p = 0.05, 0.50, 0.95$ as well as $n = 10$ and 25 only. The patterns we observe here for RB and RMSE of the estimators hold for other values of p and n as well.

Figures 15 - 20 present the plots of RB and RMSE of the three estimators of ξ_p for $p = 0.05$ (a small value), 0.50 (the median value), 0.95 (a high value) and $n = 10$ (small), 25 (moderately large).

Remark 5 The observed trends of our comprehensive simulation study to compare $\widehat{\xi}_{p(MME)}$, $\widehat{\xi}_{p(MLE)}$ and $\widehat{\xi}_{p(ORE)}$ have been summarised as follows.

(a) When n is "small" ($n < 25$) and p is "small" ($p \leq 0.10$), in terms of RB, all the three estimators are positively biased, and $\widehat{\xi}_{p(ORE)}$ has the best performance (i.e., least biased), followed by $\widehat{\xi}_{p(MME)}$ and $\widehat{\xi}_{p(MLE)}$ which are close to each other. Exactly the same pattern has been observed in terms of RMSE also.

(b) When n is "small" ($n < 25$) and p is "median" (i.e., $p = 0.50$), $\widehat{\xi}_{p(ORE)}$ still has the smallest absolute RB, though it is negative for small values of δ ($\delta < 1.00$). In terms of RMSE, $\widehat{\xi}_{p(ORE)}$ no longer enjoys the best performance for all δ. It has been noted that when $\delta < 1$, $\widehat{\xi}_{p(ORE)}$ is the best, but for $\delta > 1$ both $\widehat{\xi}_{p(MLE)}$ and $\widehat{\xi}_{p(MME)}$ over take $\widehat{\xi}_{p(ORE)}$.

(c) When n is "small" ($n < 25$) and p is "large" (i.e., $p \geq 0.90$), then in terms of RB, $\widehat{\xi}_{p(ORE)}$ has a high negative value, and $\widehat{\xi}_{p(MLE)}$ as well as $\widehat{\xi}_{p(MME)}$, which are close to each other, tend to be better than $\widehat{\xi}_{p(ORE)}$. Also, in terms of RMSE, both $\widehat{\xi}_{p(MLE)}$ and $\widehat{\xi}_{p(MME)}$ are close to each other, and better than $\widehat{\xi}_{p(ORE)}$ for $\delta > 0.50$, whereas for $\delta < 0.50$ $\widehat{\xi}_{p(ORE)}$ is preferable.

(d) When n is "large" ($n \geq 25$) and p is "small" ($p \leq 0.10$), in terms of RB, $\widehat{\xi}_{p(ORE)}$ is the best for all δ. However, in terms of RMSE, $\widehat{\xi}_{p(ORE)}$ is better than the other two only for $\delta < 1$. For $\delta > 1$, both $\widehat{\xi}_{p(MLE)}$ and $\widehat{\xi}_{p(MME)}$ are close to each other, and marginally better than $\widehat{\xi}_{p(ORE)}$.

(e) When n is "large" ($n \geq 25$) and p is "median" (i.e., $p = 0.50$), the RB patterns remain same as in (b). However, in terms of RMSE, the observed patterns are same as in (c).

(f) When n is "large" ($n \geq 25$) and p is "large" (i.e., $p \geq 0.90$), the patterns and quite different. In terms of RB, $\widehat{\xi}_{p(MLE)}$ has the overall best performance. In terms of RMSE, $\widehat{\xi}_{p(MLE)}$ is undoubtedly the best performance, closely followed by $\widehat{\xi}_{p(MME)}$ and $\widehat{\xi}_{p(ORE)}$, respectively.

If we summarize all the above findings, then for estimating ξ_p, we recommend $\widehat{\xi}_{p(ORE)}$ when p is "small", and $\widehat{\xi}_{p(MLE)}$ when p is "large". For $p \approx 0.50$, any one of these two estimators can be used.

5. Interval Estimation of the Relevant Parameters

In this section we focus our attention to interval estimation of the above-mentioned parameters, i.e., δ and σ (the model parameters), η (the mean), and ξ_p (the p^{th} percentile), based on our findings in Sections 2 - 4.

Note that the MLEs of the parameters under study have shown overall superior performance, and hence based on their asymptotic behavior we propose the following approximate $(1 - \alpha)$ level confidence interval (CI) for each of the relevant parameters.

From (9), it is noted that the asymptotic variance of $\widehat{\delta}_{MLE}$ is given as

$$AV(\widehat{\delta}_{MLE}) = [4/(n(\pi^2 - 4))](\delta^2). \tag{12}$$

Therefore, by replacing δ by its consistent estimator $\widehat{\delta}_{MLE}$ on the RHS of (12) above, we get a consistent estimator of AV($\widehat{\delta}_{MLE}$) as

$$\widehat{AV}(\widehat{\delta}_{MLE}) = [4/(n(\pi^2 - 4))](\widehat{\delta}_{MLE})^2.$$

Thus, we propose a $(1-\alpha)$ level CI for δ as

$$CI(\delta) = \widehat{\delta}_{MLE} \mp d_\star\{\widehat{AV}(\widehat{\delta}_{MLE})\}^{(1/2)},$$

where $d_\star$ is a suitable constant. We have proposed two possible values of $d_\star$, $d_\star = z_{(\alpha/2)}$ = the normal cut-off point, and $d_\star = t_{(n-1),(\alpha/2)}$ = the t cut-off point, and the resultant CIs are referred to as "z-interval" and "t-interval"

Following the same approach, our proposed $(1-\alpha)$ level CIs for σ are given as

$$CI(\sigma) = \widehat{\sigma}_{MLE} \mp d_\star\{\widehat{AV}(\widehat{\sigma}_{MLE})\}^{(1/2)},$$

where

$$\widehat{AV}(\widehat{\sigma}_{MLE}) = [\widehat{\sigma}^2_{MLE}/(n(\pi^2-4))][\pi^2/2 - 2 + (2-\ln(2)-\gamma)^2](1/\widehat{\delta}^2_{MLE}),$$

and $d_\star$ is taken as $z_{(\alpha/2)}$ as well as $t_{(n-1),(\alpha/2)}$.

Likewise, the proposed $(1-\alpha)$ level CIs for the population mean η are given as

$$CI(\eta) = \widehat{\eta}_{MLE} \mp d_\star\{\widehat{AV}(\widehat{\eta}_{MLE})\}^{(1/2)},$$

where $d_\star$ will be taken as $z_{(\alpha/2)}$ and $t_{(n-1),(\alpha/2)}$, and $\widehat{AV}(\widehat{\eta}_{MLE}) = (1/n)\widehat{AV}(\sqrt{n}\widehat{\eta}_{MLE})$ with $\widehat{AV}(\sqrt{n}\widehat{\eta}_{MLE})$ is the same expression as in (11) with δ and σ replaced by $\widehat{\delta}_{MLE}$ and $\widehat{\sigma}_{MLE}$ respectively.

For interval estimation of the p^{th} percentile, i.e., ξ_p, the expression is a bit complicated. Recall that $\widehat{\xi}_{p(MLE)}$ performs better when p is not too small. But since $\widehat{\xi}_{p(ORE)}$, which works better when p is "small", does not have any trackable asymptotic variance expression, we proceed with $\widehat{\xi}_{p(MLE)}$ only due to its available asymptotic variance formula. Write

$$\widehat{\xi}_{p(MLE)} = \widehat{\sigma}_{MLE}(c_p^{1/\widehat{\delta}_{MLE}}) = g_\star(\widehat{\delta}_{MLE}, \widehat{\sigma}_{MLE}), \text{(say)}.$$

Then using the consistency of the MLE, we have

$$\left(\begin{array}{c} \widehat{\xi}_{p(MLE)} \end{array} \right) \longrightarrow^d N\left(g_\star(\delta,\sigma), \nabla g_\star(\delta,\sigma)' I^{-1} \nabla g_\star(\delta,\sigma)\right),$$

where $g_\star(\delta,\sigma) = \sigma(c_p^{1/\delta}) = \xi_p$, and I^{-1} is given in (8). Thus, the asymptotic variance of $\widehat{\xi}_{p(MLE)}$ is given as

$$AV(\widehat{\xi}_{p(MLE)}) = \nabla g_\star(\delta,\sigma)' I^{-1} \nabla g_\star(\delta,\sigma) = (1/n)c_p^{(2/\delta)}(\sigma^2/\delta^2)K_p,$$

where $K_p = (1/(\pi^2-4))[4(\ln c_p)^2 - 4\ln c_p(2-\ln(2)-\gamma) + [\pi^2/2 - 2 + (2-\ln(2)-\gamma)^2]]$ is a known constant, details of which are given in the Appendix A.

Therefore, a consistent estimate of $AV(\widehat{\xi}_{p(MLE)})$ can be given as

$$\widehat{AV}(\widehat{\xi}_{p(MLE)}) = (1/n)c_p^{(2/\widehat{\delta}_{MLE})}(\widehat{\sigma}^2_{MLE}/\widehat{\delta}^2_{MLE})K_p.$$

Hence, the proposed $(1-\alpha)$ level CIs for ξ_p are

$$CI(\xi_p) = \widehat{\xi}_{p(MLE)} \mp d_\star\{\widehat{AV}(\widehat{\xi}_{p(MLE)})\}^{(1/2)},$$

where $d_\star$ values are chosen as above.

A thorough simulation study has been run to compare the exact coverage probability of each proposed CI when the target is $(1-\alpha)$. This has been done for various n $(= 25, 50, 75, 100$ and $200)$ and $\delta = (0.5$ to $10.0)$. W.l.g. σ has been fixed at 1. We have used $(1-\alpha) = 0.90$ and 0.95, as well as the number of replications as 2.5×10^4.

Remark 6 Let us summarize the findings of interval estimation of all the four parameters based on our comprehensive simulation study. Note that the proposed CIs are based on the asymptotic properties of the MLEs, and hence they are expected to perform well for "large" n.

(a) For the shape parameter δ estimation t-interval performs exceedingly well in terms of attaining the nominal target $(1-\alpha)$, and this happens even for "small" n. The z-interval has probability coverage (PC) very close to $(1-\alpha)$ for small n. For $n \geq 50$, both the intervals perform equally well.

(b) For σ estimation, the PC of the two CIs are far from satisfactory for $n < 50$. For $n \geq 50$, both the intervals are having PC very close to the nominal target $(1-\alpha)$.

(c) For η estimation, the performance of the two intervals are similar to those of σ estimation.

(d) For ξ_p estimation, the trends are similar to those of σ and η.

Remark 7 Since the interval estimation, as presented above, are not satisfactory for σ, η and ξ_p with "small" n, it will be taken up as a future research problem. we plan to make a 3D plot of each parameter's PC against δ and n, and then fit a suitable regression model to see if we can adjust the coefficient $d_\star$ as a function of (δ, n) to attain the nominal target $(1-\alpha)$. Then, for a given dataset, we would like to use $d_\star(\widehat{\delta}_{MLE}, n)$ to make the respective CI work. This will be studied, and reported in a future publication.

Table 1 Exact PC of the proposed CIs with $\sigma = 1$ and $(1-\alpha) = 0.90$

n	δ	$CI(\delta)$		$CI(\sigma)$		$CI(\eta)$		$CI(\xi_p)$	
		z	t	z	t	z	t	z	t
25	0.5	0.895	0.910	0.858	0.871	0.867	0.877	0.835	0.845
	1.0	0.896	0.910	0.869	0.882	0.868	0.880	0.853	0.866
	2.0	0.892	0.906	0.871	0.884	0.833	0.848	0.855	0.869
	3.0	0.892	0.907	0.871	0.884	0.816	0.832	0.862	0.875
	4.0	0.895	0.909	0.872	0.885	0.803	0.819	0.858	0.872
	5.0	0.892	0.905	0.874	0.887	0.801	0.816	0.860	0.873
	6.0	0.894	0.907	0.873	0.886	0.789	0.808	0.861	0.875
	7.0	0.893	0.908	0.877	0.891	0.789	0.807	0.862	0.875
	8.0	0.892	0.905	0.875	0.889	0.789	0.805	0.861	0.873
	9.0	0.894	0.907	0.875	0.889	0.783	0.801	0.861	0.874
	10.0	0.896	0.910	0.876	0.889	0.782	0.799	0.862	0.875
50	0.5	0.899	0.906	0.882	0.888	0.845	0.901	0.868	0.873
	1.0	0.897	0.904	0.887	0.893	0.886	0.893	0.873	0.879
	2.0	0.895	0.901	0.890	0.897	0.852	0.859	0.882	0.887
	3.0	0.897	0.904	0.887	0.894	0.827	0.835	0.881	0.888
	4.0	0.897	0.902	0.885	0.891	0.821	0.828	0.879	0.885
	5.0	0.895	0.901	0.886	0.892	0.811	0.820	0.878	0.885
	6.0	0.899	0.906	0.884	0.890	0.801	0.811	0.879	0.885
	7.0	0.896	0.903	0.887	0.893	0.799	0.807	0.879	0.886
	8.0	0.898	0.904	0.890	0.896	0.798	0.805	0.881	0.888
	9.0	0.894	0.901	0.886	0.893	0.795	0.803	0.879	0.886
	10.0	0.894	0.901	0.886	0.893	0.793	0.802	0.880	0.888

Table 1 (Continued)

n	δ	$CI(\delta)$		$CI(\sigma)$		$CI(\eta)$		$CI(\xi_p)$	
		z	t	z	t	z	t	z	t
75	0.5	0.894	0.898	0.889	0.894	0.902	0.906	0.879	0.883
	1.0	0.899	0.904	0.888	0.893	0.887	0.892	0.881	0.885
	2.0	0.900	0.905	0.891	0.895	0.854	0.859	0.887	0.893
	3.0	0.899	0.904	0.890	0.894	0.832	0.837	0.885	0.889
	4.0	0.895	0.899	0.890	0.894	0.823	0.828	0.884	0.888
	5.0	0.898	0.902	0.890	0.895	0.813	0.818	0.889	0.893
	6.0	0.897	0.901	0.892	0.896	0.808	0.813	0.885	0.889
	7.0	0.900	0.904	0.893	0.897	0.806	0.812	0.889	0.894
	8.0	0.894	0.898	0.897	0.901	0.804	0.810	0.888	0.892
	9.0	0.896	0.900	0.890	0.894	0.799	0.805	0.886	0.891
	10.0	0.900	0.904	0.893	0.896	0.796	0.802	0.888	0.892
100	0.5	0.899	0.902	0.890	0.894	0.907	0.910	0.886	0.889
	1.0	0.899	0.902	0.893	0.896	0.893	0.896	0.886	0.889
	2.0	0.899	0.901	0.895	0.898	0.856	0.860	0.889	0.892
	3.0	0.894	0.898	0.892	0.895	0.834	0.838	0.891	0.894
	4.0	0.898	0.901	0.895	0.898	0.825	0.829	0.891	0.895
	5.0	0.897	0.900	0.893	0.897	0.818	0.822	0.887	0.890
	6.0	0.899	0.903	0.894	0.896	0.811	0.816	0.887	0.890
	7.0	0.900	0.903	0.895	0.898	0.808	0.813	0.891	0.894
	8.0	0.899	0.902	0.896	0.900	0.804	0.808	0.890	0.893
	9.0	0.897	0.900	0.894	0.897	0.802	0.807	0.890	0.893
	10.0	0.901	0.904	0.893	0.896	0.793	0.797	0.890	0.893
200	0.5	0.899	0.901	0.896	0.898	0.914	0.916	0.890	0.892
	1.0	0.897	0.898	0.894	0.896	0.895	0.897	0.894	0.896
	2.0	0.899	0.900	0.896	0.898	0.860	0.862	0.892	0.893
	3.0	0.901	0.902	0.897	0.898	0.840	0.842	0.896	0.898
	4.0	0.899	0.901	0.897	0.890	0.827	0.829	0.896	0.898
	5.0	0.900	0.901	0.900	0.902	0.820	0.822	0.896	0.898
	6.0	0.900	0.901	0.899	0.900	0.815	0.817	0.894	0.896
	7.0	0.901	0.903	0.894	0.896	0.808	0.811	0.894	0.896
	8.0	0.895	0.897	0.897	0.898	0.805	0.808	0.894	0.895
	9.0	0.902	0.903	0.898	0.900	0.806	0.808	0.893	0.895
	10.0	0.898	0.900	0.897	0.899	0.803	0.806	0.895	0.896

6. Applications of 2P-GHND for two real-life datasets

In this section, we consider two real-life datasets, and demonstrate the usage of the three estimation techniques to fit a 2P-GHND model to the datasets. The goodness of fit of an estimated model to a dataset is measured by the Kolmogorov-Smirnov (KS) distance defined as

$$D^{KS} = \sup_{x\in(0,\infty)} |F_n(x) - \widehat{F}(x)|,$$

where $F_n(x)$ is the empirical cdf, i.e., $F_n(x)$=[(Number of X_i's $\leq x$)/ n] and $\widehat{F}(x)$ is a fitted cdf, i.e., $\widehat{F}(x) = 2\Phi[(x/\widehat{\sigma})^{\widehat{\delta}}] - 1$, where $(\widehat{\delta}, \widehat{\sigma})$ can be one of the three estimators discussed earlier.

After identifying the most suitable estimator to fit the 2P-GHND model, we compare the fitted model of 2P-GHND with a few other commonly used models, such as, gamma, lognormal and

Table 2 Exact PC of the proposed CIs with $\sigma = 1$ and $(1 - \alpha) = 0.95$

n	δ	$CI(\delta)$		$CI(\sigma)$		$CI(\eta)$		$CI(\xi_p)$	
		z	t	z	t	z	t	z	t
25	0.5	0.947	0.959	0.912	0.924	0.911	0.921	0.870	0.881
	1.0	0.950	0.962	0.921	0.935	0.920	0.933	0.891	0.902
	2.0	0.949	0.960	0.924	0.937	0.899	0.914	0.902	0.912
	3.0	0.948	0.960	0.927	0.939	0.886	0.902	0.905	0.917
	4.0	0.947	0.959	0.926	0.939	0.873	0.891	0.905	0.916
	5.0	0.947	0.959	0.930	0.942	0.872	0.889	0.906	0.917
	6.0	0.948	0.959	0.928	0.940	0.861	0.880	0.904	0.916
	7.0	0.947	0.960	0.925	0.939	0.858	0.876	0.905	0.917
	8.0	0.948	0.960	0.928	0.941	0.857	0.876	0.904	0.917
	9.0	0.947	0.959	0.926	0.938	0.854	0.872	0.906	0.918
	10.0	0.945	0.956	0.929	0.941	0.854	0.873	0.904	0.916
50	0.5	0.948	0.954	0.931	0.936	0.935	0.940	0.906	0.911
	1.0	0.947	0.953	0.935	0.940	0.933	0.939	0.916	0.920
	2.0	0.951	0.956	0.938	0.944	0.910	0.918	0.924	0.929
	3.0	0.948	0.954	0.940	0.946	0.896	0.904	0.924	0.930
	4.0	0.948	0.954	0.939	0.946	0.885	0.894	0.928	0.934
	5.0	0.949	0.955	0.940	0.945	0.879	0.887	0.930	0.935
	6.0	0.952	0.958	0.939	0.945	0.877	0.887	0.927	0.933
	7.0	0.949	0.955	0.940	0.947	0.869	0.878	0.930	0.935
	8.0	0.948	0.953	0.939	0.945	0.869	0.877	0.929	0.935
	9.0	0.947	0.954	0.942	0.948	0.869	0.879	0.929	0.934
	10.0	0.947	0.953	0.939	0.945	0.864	0.873	0.926	0.932
75	0.5	0.950	0.954	0.937	0.940	0.943	0.946	0.921	0.924
	1.0	0.951	0.954	0.942	0.945	0.941	0.945	0.929	0.932
	2.0	0.946	0.950	0.942	0.946	0.915	0.920	0.928	0.931
	3.0	0.949	0.953	0.944	0.948	0.901	0.906	0.933	0.937
	4.0	0.950	0.953	0.943	0.946	0.887	0.893	0.934	0.938
	5.0	0.948	0.952	0.939	0.943	0.881	0.888	0.933	0.937
	6.0	0.952	0.955	0.942	0.946	0.881	0.887	0.936	0.940
	7.0	0.947	0.952	0.942	0.945	0.877	0.884	0.935	0.939
	8.0	0.947	0.950	0.941	0.945	0.870	0.876	0.932	0.936
	9.0	0.948	0.953	0.944	0.947	0.874	0.880	0.937	0.941
	10.0	0.948	0.952	0.943	0.948	0.864	0.871	0.937	0.941

Weibull by using the Akaike information criterion (AIC). The formula of AIC is defined as

$$AIC = -2L_* + 2k,$$

where L_* is the log-likelihood function of the distribution with estimated parameters and k is a number of a parameters.

Example 1 We use the data on the recored failure time (RFT) (given in Tanaka et al. (2018)) for a small sample case ($n = 10$). We fit a 2P-GHND model and estimate the parameters δ and σ by using MME, MLE and ORE as shown in Table 3. After that, we show the empirical relative frequency histogram and the fitted 2P-GHND models in Figure 21. They all look positively skewed, and the MME as well as the MLE provide much better fit than the ORE as shown in Table 4. Note that the

Table 2 (Continued)

n	δ	$CI(\delta)$		$CI(\sigma)$		$CI(\eta)$		$CI(\xi_p)$	
		z	t	z	t	z	t	z	t
100	0.5	0.951	0.954	0.939	0.942	0.947	0.949	0.926	0.929
	1.0	0.948	0.952	0.944	0.947	0.944	0.946	0.933	0.936
	2.0	0.948	0.952	0.944	0.947	0.916	0.920	0.937	0.940
	3.0	0.948	0.951	0.943	0.945	0.900	0.904	0.938	0.940
	4.0	0.949	0.952	0.945	0.948	0.894	0.898	0.936	0.939
	5.0	0.948	0.951	0.947	0.949	0.887	0.891	0.937	0.940
	6.0	0.951	0.953	0.946	0.949	0.883	0.888	0.938	0.941
	7.0	0.951	0.953	0.946	0.949	0.880	0.885	0.939	0.941
	8.0	0.951	0.954	0.945	0.948	0.874	0.879	0.938	0.941
	9.0	0.952	0.954	0.946	0.949	0.875	0.880	0.938	0.941
	10.0	0.951	0.954	0.945	0.948	0.870	0.874	0.938	0.941
200	0.5	0.949	0.950	0.944	0.945	0.954	0.956	0.939	0.941
	1.0	0.948	0.949	0.947	0.949	0.947	0.949	0.944	0.946
	2.0	0.951	0.952	0.949	0.950	0.922	0.924	0.944	0.945
	3.0	0.950	0.952	0.947	0.948	0.903	0.905	0.942	0.944
	4.0	0.950	0.951	0.949	0.950	0.896	0.898	0.947	0.948
	5.0	0.950	0.951	0.947	0.948	0.889	0.891	0.945	0.946
	6.0	0.951	0.952	0.948	0.950	0.887	0.889	0.944	0.946
	7.0	0.947	0.949	0.946	0.947	0.878	0.880	0.944	0.945
	8.0	0.952	0.953	0.947	0.949	0.878	0.880	0.944	0.946
	9.0	0.950	0.951	0.949	0.950	0.877	0.879	0.944	0.945
	10.0	0.950	0.952	0.948	0.949	0.874	0.876	0.945	0.946

performance of model fitting by MME and MLE are almost identical.

Next, we compare our best fitted 2P-GHND (using the MLE to maintain uniformity in all estimations) with three other popular models, namely - gamma, Weibull and lognormal, and the goodness of fit has been measured by the AIC as reported in Table 5.

Remark 8 Figure 22 gives the empirical cdf (the step function) along with the three fitted 2P-GHND cdfs. Obviously, the ORE fitted cdf appears to be over estimating the cdf (since it is mostly higher than the empirical cdf), whereas the MLE and the MME fitted cdfs (which are indistinguishable) appear to be matching the empirical cdf well. Also, the Table 5 confirms the 2P-GHND is the best fit for the RFT data.

Table 3 Estimates of parameter δ and σ for RFT data

$\widehat{\delta}$	$\widehat{\sigma}$
$\widehat{\delta}_{MME} = 1.859$	$\widehat{\sigma}_{MME} = 957.802$
$\widehat{\delta}_{MLE} = 1.933$	$\widehat{\sigma}_{MLE} = 959.737$
$\widehat{\delta}_{ORE} = 1.723$	$\widehat{\sigma}_{ORE} = 836.755$

Example 2 This example deals with the monthly total rainfall (MTR) (in mm) on the east coast of Australia in the State of New South Wales as given by Gómez and Vidal (2016) for a large sample

Table 4 The KS distance for three fitted models for the Example 1

Method	MME	MLE	ORE
D^{KS}	0.075	0.077	0.187

Table 5 The AIC of the fitted models for the Example 1

Model	Estimated Parameter		AIC
2P-GHND	$\widehat{\delta} = 1.933$	$\widehat{\sigma} = 959.737$	148.956
Gamma	$\widehat{\alpha} = 3.939$	$\widehat{\beta} = 198.501$	150.103
Weibull	$\widehat{\alpha} = 882.818$	$\widehat{\beta} = 2.390$	149.283
Log-normal	$\widehat{\mu} = 6.529$	$\widehat{\sigma} = 0.559$	151.347

with $n = 83$ (the total monthly rainfall in the study area appears to be iid over each year). We used these data to obtain the δ and σ estimates for a fitted 2P-GHND as shown in Table 6. After that, we show the empirical relative frequency histogram and the fitted 2P-GHND models in Figure 23. They all look positively skewed, and the performance of model fitting by MME and MLE are almost identical.

Table 6 Estimates of parameter δ and σ for MTR

$\widehat{\delta}$	$\widehat{\sigma}$
$\widehat{\delta}_{MME} = 1.106$	$\widehat{\sigma}_{MME} = 42.635$
$\widehat{\delta}_{MLE} = 1.108$	$\widehat{\sigma}_{MME} = 42.635$
$\widehat{\delta}_{ORE} = 1.027$	$\widehat{\sigma}_{ORE} = 41.799$

Table 7 The KS distance for three fitted models for the Example 2

Method	MME	MLE	ORE
D^{KS}	0.062	0.063	0.067

Table 8 The AIC of the fitted models for the Example 2

Model	Estimated Parameter		AIC
2P-GHND	$\widehat{\delta} = 1.108$	$\widehat{\sigma} = 42.635$	741.021
Gamma	$\widehat{\alpha} = 1.516$	$\widehat{\beta} = 22.384$	747.309
Weibull	$\widehat{\alpha} = 36.912$	$\widehat{\beta} = 1.366$	744.489
Log-normal	$\widehat{\mu} = 3.159$	$\widehat{\sigma} = 1.019$	767.192

Remark 9 Figure 24 shows that three estimation methods provide nearly similar 2P-GHND cdfs, perhaps due to a large n. Table 7 confirms that three estimation methods are close to each other. Also, Table 8 shows that the 2P-GHND appears to provide a better fit than the other three models.

7. Conclusions

This work presents a comprehensive study of three estimation methods for various parameters related to a 2P-GHND which can be beneficial to fit real-life datasets. It is not claimed that the 2P-GHND is going to be the best model for all datasets. However, as our two applications show, the 2P-GHND should be included along with other popular positively skewed models (such as - Gamma, Weibull and Log-normal) in model fitting to see which one can produce the best fit, and then only subsequent inferences can be drawn. If 2P-GHND comes out as the best fitted model, then we can estimate its percentiles as presented in Section 4, since percentiles are of major interest in applied problems.

Acknowledgements

We sincerely thank the two anonymous referees who provided valuable comments and suggestions which have helped us immensely in improving the presentation of the material of this paper.

References

Cooray K, Ananda MMA. A generalized of the half-normal distribution with applications to lifetime data. Commun Stat Theory Methods. 2008; 37: 1323-1337.

Efron B. The Jackknife, the Bootstrap and Other Resampling Plans. C.A.: Duxbury, Thomson Learning; 1982.

Gómez YM, Vidal I. A generalization of the half-normal distribution. Appl Math J Chinese Univ. 2016; 31(4): 409-424.

Mazucheli J, Dey S. Bias-corrected maximum likelihood estimation of the parameters of the generalized half-normal distribution. J Stat Comput Sim. 2018; 88(6): 1027-1038.

Pewsey A. Large sample inference for the general half-normal distribution. Comm Statist Theory Methods. 2002; 31(7):1045-1054.

Pewsey A. Improved likelihood based inference for the general half normal distribution. Commun Statist Theory Methods. 2004; 33(2): 197-204.

Stacy EW. A generalization of gamma distribution. Institute of Mathematical Statistics. 1962; 33(3): 1187-1192.

Stacy EW, Mihram GA. Parameter estimation for a generalization gamma distribution. Technometrics. 1965; 7(3): 349-358.

Tanaka H, Pal N, Lim WK. On improved estimation under Weibull model. J Stat Theory Pract. 2018; 12(1): 48-65.

Appendix A: The asymptotic variance of the $\widehat{\xi}_{p(MLE)}$

This part shows the details how to estimated the $AV(\widehat{\xi}_{p(MLE)})$. From

$$\widehat{\xi}_p = \widehat{\sigma}\left\{\Phi^{-1}\left(\frac{p+1}{2}\right)\right\}^{1/\widehat{\delta}} = \widehat{\sigma}h(\widehat{\delta}) = g_\star(\widehat{\delta},\widehat{\sigma}),$$

and

$$\left(\ \widehat{\xi}_p\ \right) \longrightarrow^d N\left(g_\star(\delta,\sigma), \nabla g_\star(\delta,\sigma)' I^{-1} \nabla g_\star(\delta,\sigma)\right),$$

where

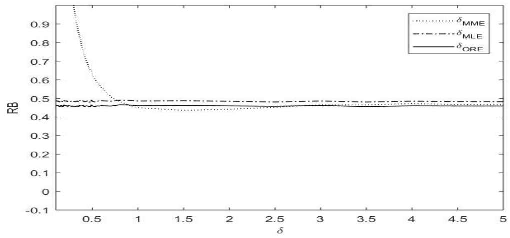
$$\nabla g_\star(\delta,\sigma) = \xi_p = \begin{bmatrix}(\partial/\partial\delta)g_\star(\delta,\sigma)\\ \\ (\partial/\partial\sigma)g_\star(\delta,\sigma)\end{bmatrix} = \begin{bmatrix}\sigma(-\ln c_p/\delta^2)c_p^{(1/\delta)}\\ \\ c_p^{(1/\delta)}\end{bmatrix} = c_p^{(1/\delta)}\begin{bmatrix}\sigma(-\ln c_p/\delta^2)\\ \\ 1\end{bmatrix},$$

where $c_p = \Phi^{-1}((p+1)/2)$. So,

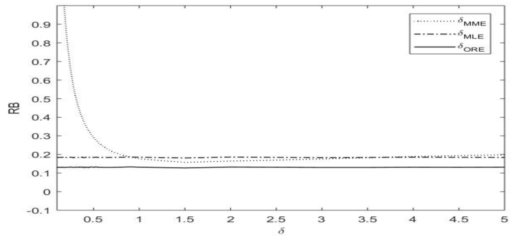
$$\begin{aligned}\nabla g_\star(\delta,\sigma)' I^{-1} \nabla g_\star(\delta,\sigma) &= c_p^{(2/\delta)} \begin{bmatrix} \sigma(-\ln c_p/\delta^2) \\ 1 \end{bmatrix}' I^{-1} \begin{bmatrix} \sigma(-\ln c_p/\delta^2) \\ 1 \end{bmatrix} \\ &= (1/n) c_p^{(2/\delta)} (\sigma^2/\delta^2) K_p,\end{aligned}$$

where $K_p = (1/(\pi^2-4))[4(\ln c_p)^2 - 4\ln c_p(2 - ln(2) - \gamma) + [\pi^2/2 - 2 + (2 - ln(2) - \gamma)^2]]$.

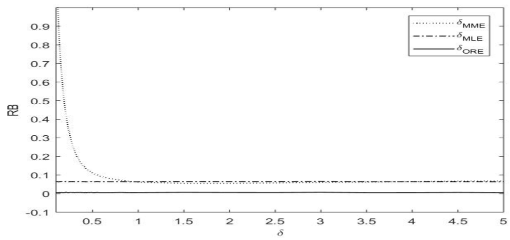
Appendix B: Figures for estimating model parameters δ and σ of 2P-GHND in term of RB and RMSE



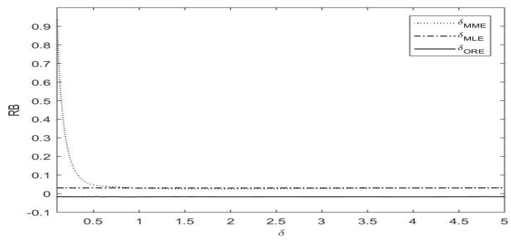
(a) $n = 5$



(b) $n = 10$

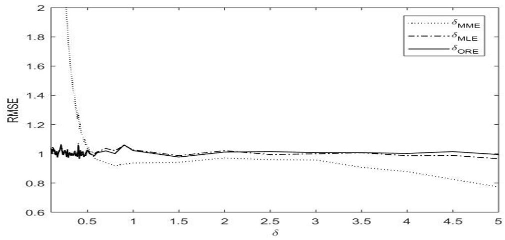


(c) $n = 25$

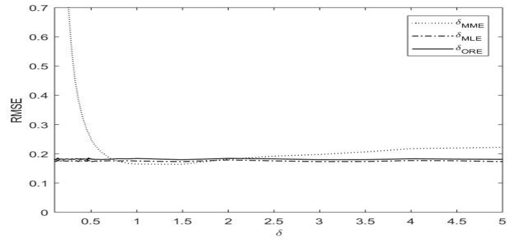


(d) $n = 50$

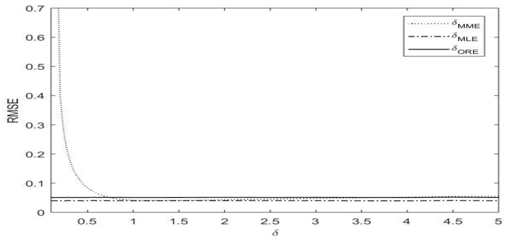
Figure 2 RB of $\widehat{\delta}_{MME}$, $\widehat{\delta}_{MLE}$ and $\widehat{\delta}_{ORE}$



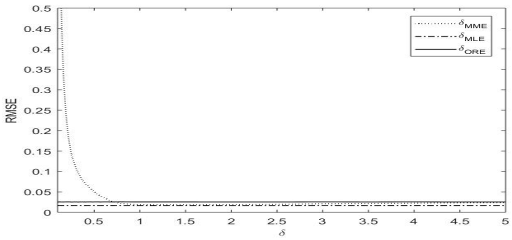
(a) $n = 5$



(b) $n = 10$

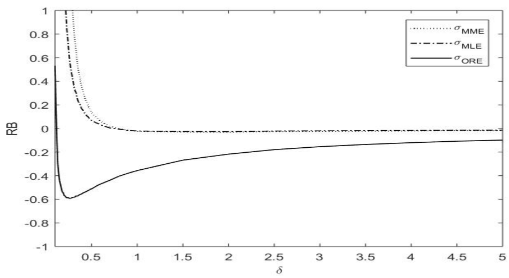


(c) $n = 25$

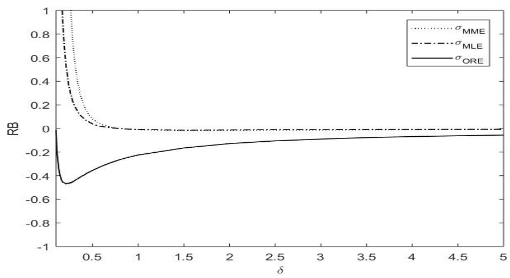


(d) $n = 50$

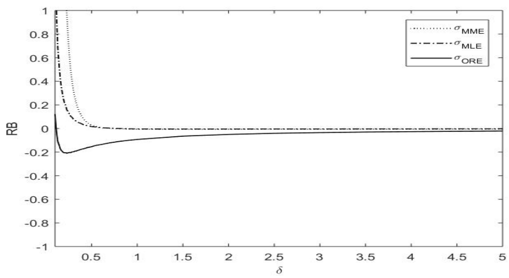
Figure 3 RMSE of $\widehat{\delta}_{MME}$, $\widehat{\delta}_{MLE}$ and $\widehat{\delta}_{ORE}$



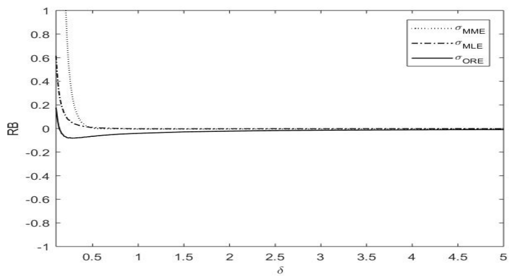
(a) $n = 5$



(b) $n = 10$

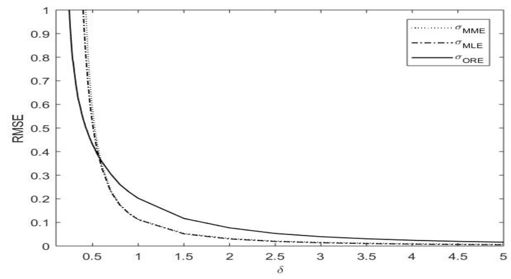


(c) $n = 25$

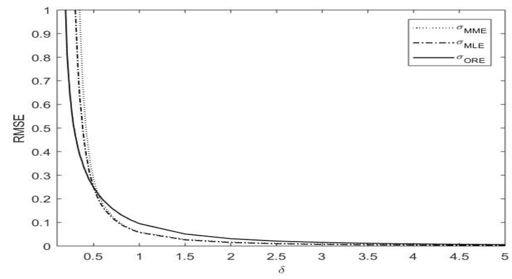


(d) $n = 50$

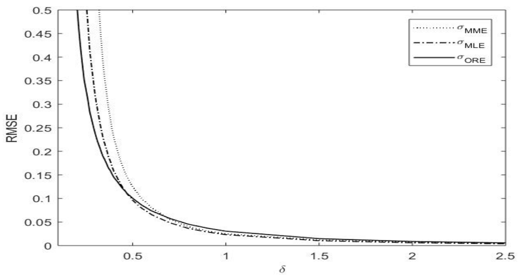
Figure 4 RB of $\widehat{\sigma}_{MME}$, $\widehat{\sigma}_{MLE}$ and $\widehat{\sigma}_{ORE}$



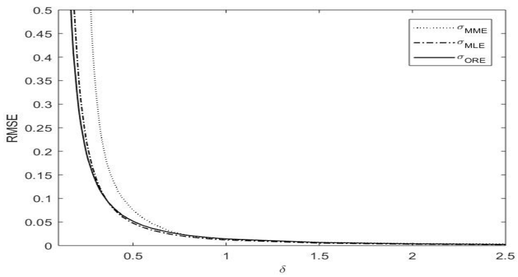
(a) $n = 5$



(b) $n = 10$



(c) $n = 25$



(d) $n = 50$

Figure 5 RMSE of $\widehat{\sigma}_{MME}$, $\widehat{\sigma}_{MLE}$ and $\widehat{\sigma}_{ORE}$

Appendix C: Figures of the asymptotic behaviour of δ and σ

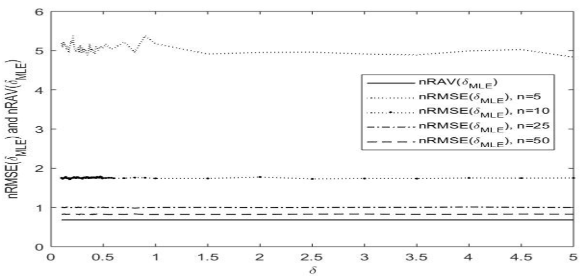


Figure 6 Comparison between RMSE and RAV of $\widehat{\delta}_{MLE}$

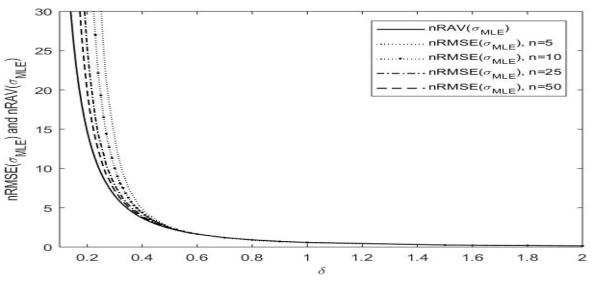
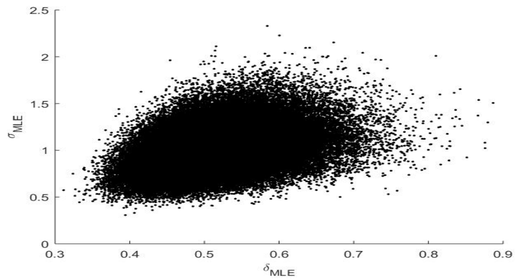
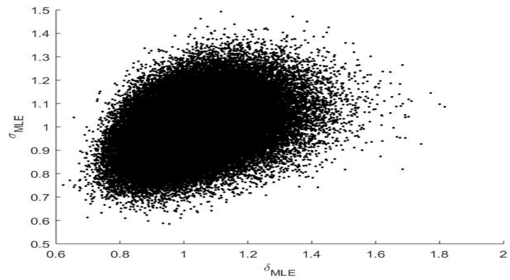


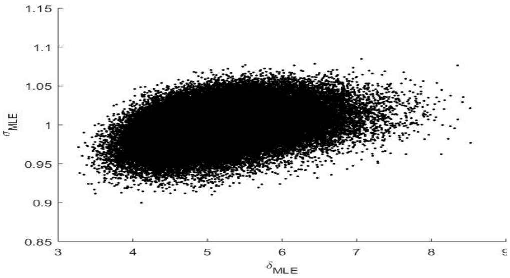
Figure 7 Comparison between RMSE and RAV of $\widehat{\sigma}_{MLE}$



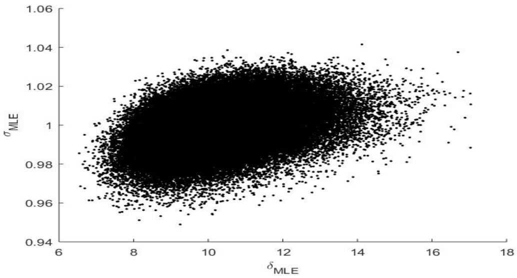
(a) $\delta = 0.5, n = 50$



(b) $\delta = 1, n = 50$

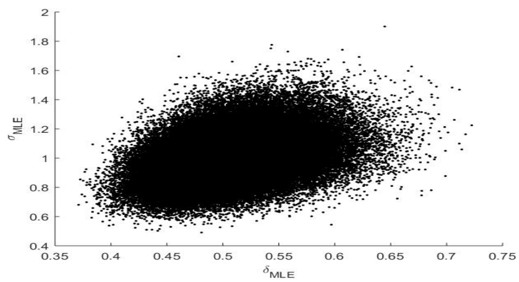


(c) $\delta = 5, n = 50$

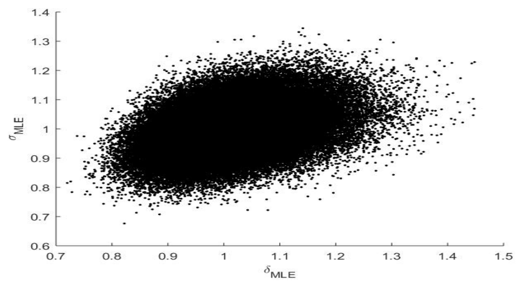


(d) $\delta = 10, n = 50$

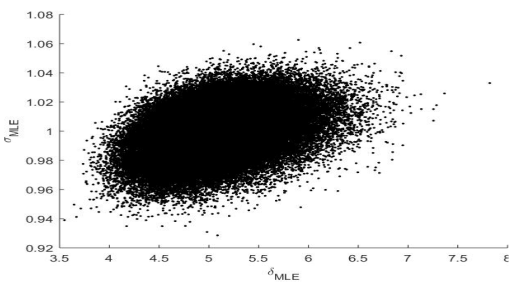
Figure 8 Asymptotic Correlation Between $\widehat{\delta}_{MLE}$ and $\widehat{\sigma}_{MLE}$ for $n = 50$



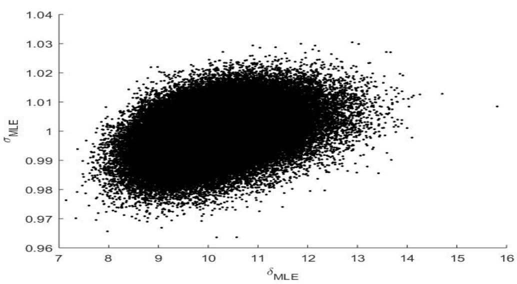
(a) $\delta = 0.5, n = 100$



(b) $\delta = 1, n = 100$



(c) $\delta = 5, n = 100$



(d) $\delta = 10, n = 100$

Figure 9 Asymptotic Correlation Between $\widehat{\delta}_{MLE}$ and $\widehat{\sigma}_{MLE}$ for $n = 100$

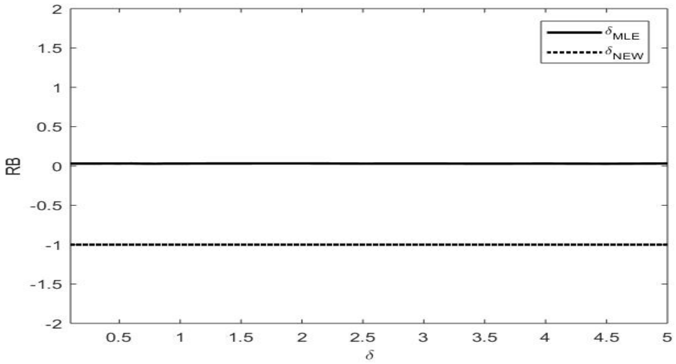


Figure 10 RB of $\widehat{\delta}_{MLE}$ and $\widehat{\delta}_{NEW}$ for $n = 50$

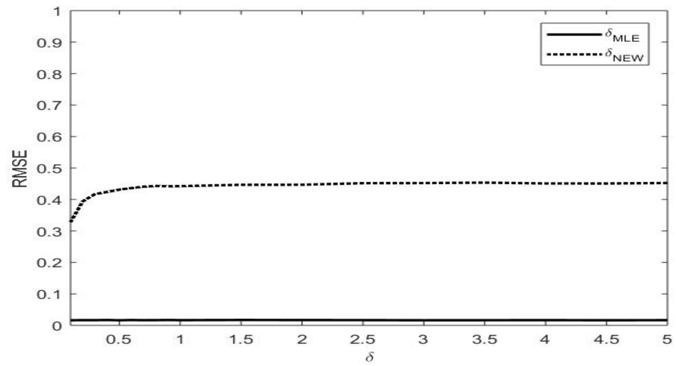


Figure 11 RMSE of $\widehat{\delta}_{MLE}$ and $\widehat{\delta}_{NEW}$ for $n = 50$

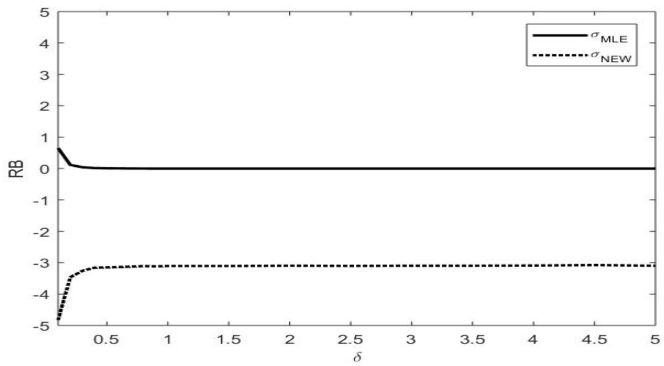


Figure 12 RB of $\widehat{\sigma}_{MLE}$ and $\widehat{\sigma}_{NEW}$ for $n = 50$

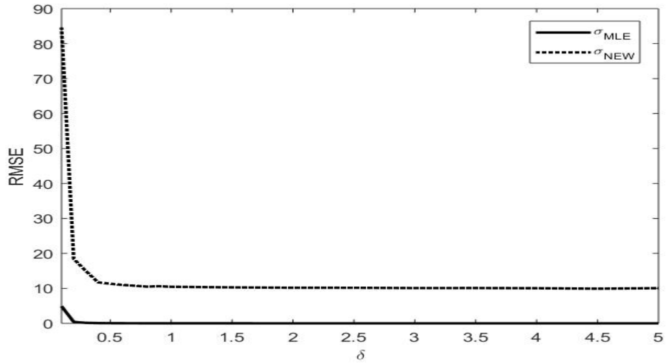


Figure 13 RMSE of $\widehat{\sigma}_{MLE}$ and $\widehat{\sigma}_{NEW}$ for $n = 50$

Appendix D: Figures of the asymptotic behaviour of the model mean (η)

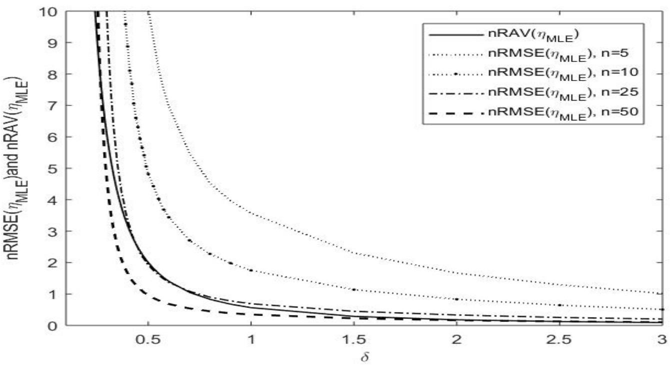
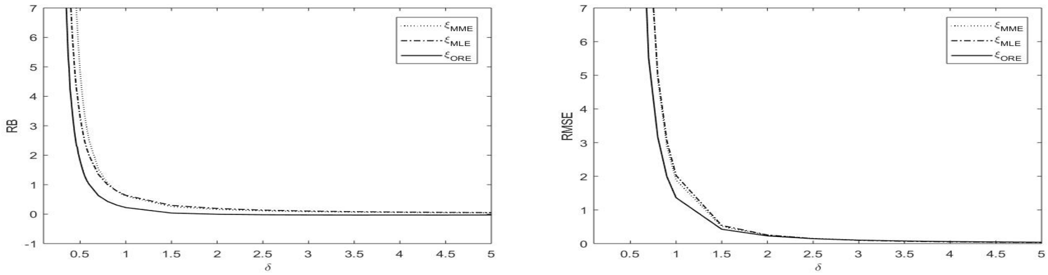


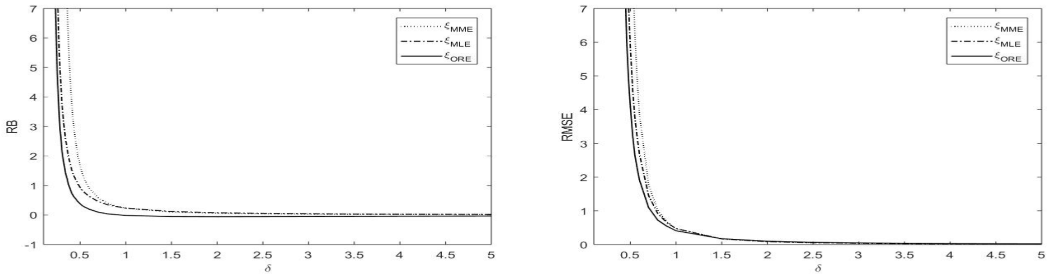
Figure 14 Comparison between RMSE and RAV of $\widehat{\eta}_{MLE}$

Appendix E: Figures for comparing of the model p^{th} percentile in term of RB and RMSE



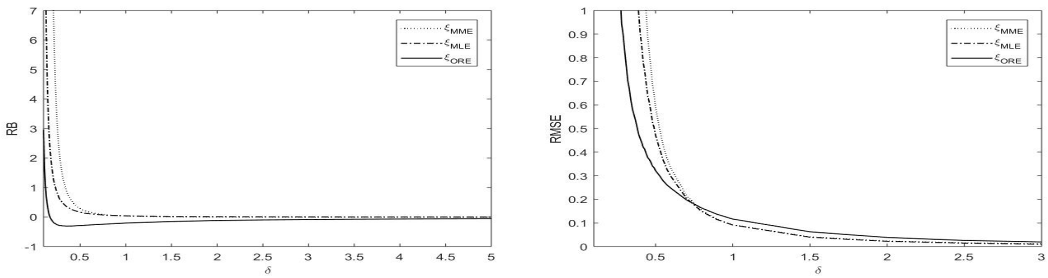
(a) **(b)**

Figure 15 RB (in (a)) and RMSE (in (b)) of $\widehat{\xi}_p$ for $n = 10$, $p = 0.05$



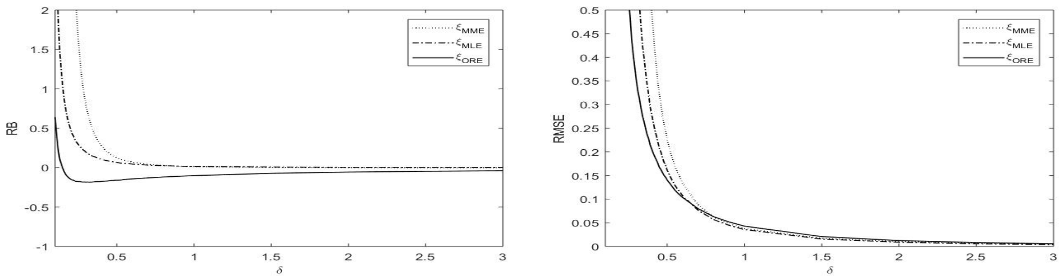
(a) **(b)**

Figure 16 RB (in (a)) and RMSE (in (b)) of $\widehat{\xi}_p$ for $n = 25$, $p = 0.05$



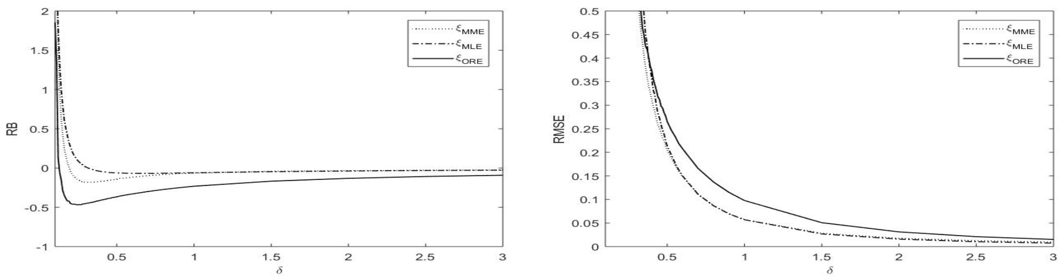
(a) **(b)**

Figure 17 RB (in (a)) and RMSE (in (b)) of $\widehat{\xi}_p$ for $n = 10$, $p = 0.50$



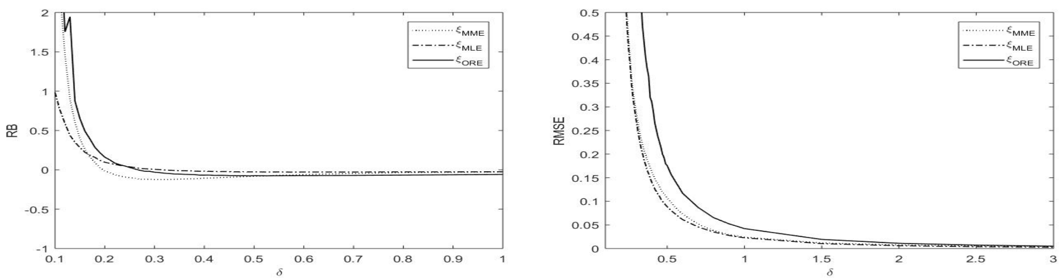
(a) **(b)**

Figure 18 RB (in (a)) and RMSE (in (b)) of $\widehat{\xi_p}$ for $n = 25$, $p = 0.50$



(a) **(b)**

Figure 19 RB (in (a)) and RMSE (in (b)) of $\widehat{\xi_p}$ for $n = 10$, $p = 0.95$



(a) **(b)**

Figure 20 RB (in (a)) and RMSE (in (b)) of $\widehat{\xi_p}$ for $n = 25$, $p = 0.95$

Appendix F: Figures for Example 1

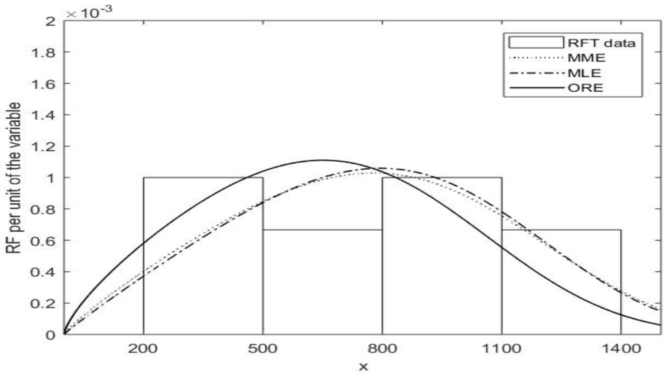


Figure 21 Empirical relative frequency histogram for RFT data along with three fitted 2P-GHND pdf curves

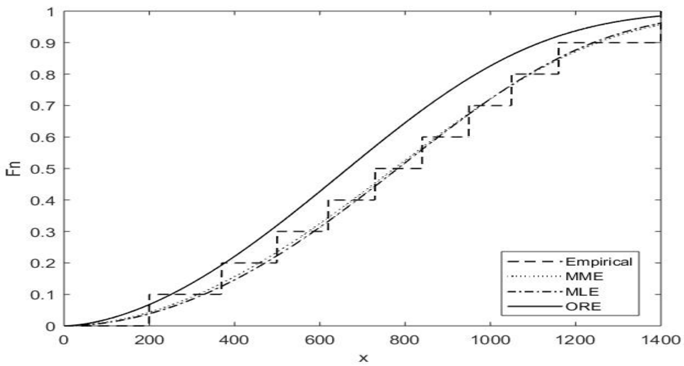


Figure 22 Empirical cdf for RFT data along with three fitted 2P-GHND cdf curves

Appendix G: Figures for Example 2

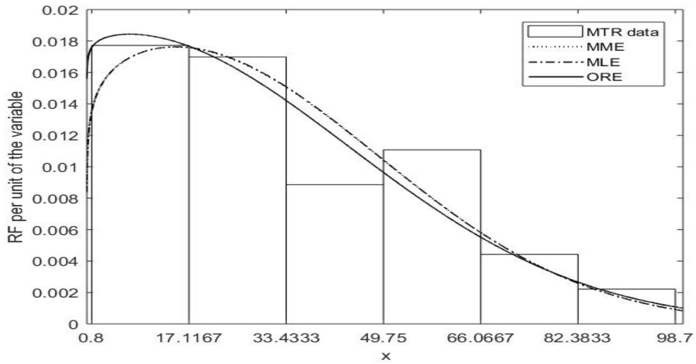


Figure 23 Empirical relative frequency histogram for MTR data along with three fitted 2P-GHND pdf curves

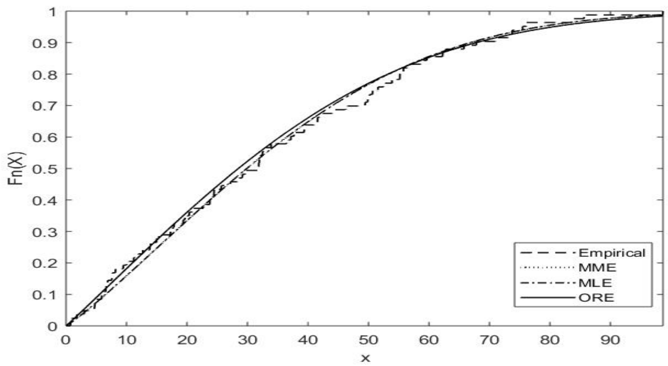


Figure 24 Empirical cdf for MTR data along with three fitted 2P-GHND cdf curves

Appendix H: The codes for the data set in Example 1

$w = [20037050062073084095010501 1601400]$
$n = length(w); l = 1; sigma = 1;$
$forl = 1 : n; c(l) = norminv(0.5 * (1 + l/n), 0, 1); end$
$c(n) = 4; lnc = log(c); lncavg = sum(lnc)/n;$
$x0 = random('Uniform', 0.01, 20, 1, 1);$
$sumx = sum(w); sumxx = sum(w. * w);$
$x = fsolve('MME', x0, optimset('fsolve'), sumx, sumxx, n);$
$whilex < 0$
$x1 = random('Uniform', 0.01, 20, 1, 1);$
$x = fsolve('MME', x1, optimset('fsolve'), sumx, sumxx, n); end$
$sigmaMME = [sumx./n]./[sqrt((2.^(1./x))./pi) * (gamma((x + 1)./(2. * x)))];$
$deltaMMEhat = x;$
$sigmaMMEhat = sigmaMME;$
$sumlogx = sum(log(w));$
$y = fsolve('MLE', x0, optimset('fsolve'), w, sumlogx, n);$
$whiley < 0$
$x1 = random('Uniform', 0.01, 20, 1, 1);$
$y = fsolve('MLE', x1, optimset('fsolve'), w, sumlogx, n);$
end
$sigmaMLE = [sum(w.^(2. * y))./n].^(1./(2 * y));$
$deltaMLEhat = y;$
$sigmaMLEhat = sigmaMLE;$
$MeanMLE = sigmaMLE. * (sqrt((2.^(1./y))/pi)). * (gamma((1 + y)./(2. * y)));$
$s = sort(w); r = log(s); ravg = sum(r)/n;$
$delta0 = sum((r - ravg). * (lnc - lncavg))./sum((lnc - lncavg).^2);$
$sigma0 = ravg - (delta0 * lncavg);$
$deltaOREhat = 1./delta0;$
$sigmaOREhat = exp(sigma0);$