Mathematics and Statistics 11(3): 516-527, 2023 http://www.hrpub.org
DOI: 10.13189/ms.2023.110307

Maximum Likelihood Estimation of the Weighted Mixture Generalized Gamma Distribution

Wikanda Phaphan[1,2], Teerawat Simmachan[3,4,*], Ibrahim Abdullahi[5]

[1]Department of Applied Statistics, Faculty of Applied Science, King Mongkut's University of Technology North Bangkok, Bangkok, Thailand

[2]Research Group in Statistical Learning and Inference, KMUTNB, Bangkok 10800, Thailand

[3]Department of Mathematics and Statistics, Faculty of Science and Technology, Thammasat University, Pathum Thani, Thailand

[4]Thammasat University Research Unit in Data Learning, Thammasat University, Pathum Thani, Thailand

[5]Department of Mathematics and Statistics, Faculty of Science, Yobe State University, Damaturu, Nigeria

Received December 23, 2022; Revised March 21, 2023; Accepted April 7, 2023

Cite This Paper in the following Citation Styles

***(a)**: [1] Wikanda Phaphan, Teerawat Simmachan, Ibrahim Abdullahi, "Maximum Likelihood Estimation of the Weighted Mixture Generalized Gamma Distribution," Mathematics and Statistics, Vol.11, No.3, pp. 516-527, 2023. DOI: 10.13189/ms.2023.110307*

***(b)**: Wikanda Phaphan, Teerawat Simmachan, Ibrahim Abdullahi, (2023). Maximum Likelihood Estimation of the Weighted Mixture Generalized Gamma Distribution. Mathematics and Statistics, 11(3), 516-527. DOI: 10.13189/ms.2023.110307*

Abstract The three-parameter weighted mixture generalized gamma (WMGG) distribution was developed from the four-parameter mixture generalized gamma (MGG) distribution since the parameter estimation of MGG distribution faced with the problem. The estimate of the weighted parameter p was out of the interval $[0, 1]$. The previous study proposed the maximum likelihood estimators (MLEs) of the WMGG distribution. However, their MLEs were written in nonlinear equations, and certain iterative methods were necessarily needed to solve numerically. The three parameters λ, β, and α were estimated by the quasi-Newton method. Nevertheless, this method performed well only the parameter λ. This motivated the main objective of this work. Consequently, the parameter estimation of the WMGG was further improved. This article developed two maximum likelihood estimation methods: expectation-maximization (EM) algorithm and simulated annealing algorithm of the three parameters of the WMGG distribution. These two methods were compared to the previous study's quasi-Newton method. Monte Carlo simulation technique was employed to assess the algorithm's performance. Sample sizes ranged from small to large as 10, 30, 50, and 100. The simulation was repeated 10,000 rounds in each scenario. Assessment criteria were the mean square error (MSE) and bias. The results revealed that the EM algorithm outperformed the other methods. Furthermore, the quasi-Newton method had the lowest efficiency.

Keywords Mixture Generalized Gamma Distribution, Expectation-Maximization Algorithm, Simulated Annealing, Quasi-Newton Method

1 Introduction

In survival analysis, the survival time T is a sum of the n survival times T_i in each state. Assume that failure or death occurs at regular intervals. In the initial state, the failure or death occurs far away from the starting time with the period of time T_1. In the second state, the failure happens after the beginning of the state within the period of time T_2. The process arises continuously until the failure ends at the final state with the time T_n. Suppose that the survival times T_i in each state are independent and identically exponentially distributed. The probability density function or PDF of the variable T_i is defined as

$$f(t_i) = \lambda e^{-\lambda t_i}, \tag{1}$$

where $\lambda > 0$, $t_i > 0$ and $i = 1, 2, ..., n$. The distribution of $T = T_1+T_2+...+T_n$ is called the Erlang distribution proposed by Erlang [1]. The distribution's first application was a study of waiting times in telephone conversations. His concept has been widely used in queuing theory and survival analysis.

A generalization form of the Erlang distribution is established when we replace the parameter n, which is an integer of $1, 2, ...$, with the parameter γ by $\gamma \in \mathbf{R}^+$. As a result, T has a gamma distribution with a shape parameter γ. With a discretized shape parameter, therefore, the Erlang distribution is a particular case of the gamma distribution. The gamma distribution is one of the most popular families of continuous probability distributions due to its flexible property.

As previously mentioned, there are numerous contributions in the literature concerning the gamma distribution. One of those contributions involves incorporating the gamma distribution into a family of mixed distributions. For instance, the generalized gamma distribution also known as the three-parameter gamma distribution was originally proposed by [2]. This distribution contains well known special cases such as Weibull [3], log-normal [4], gamma [5], Rayleigh [6] distributions, and other distributions. The generalized gamma distribution includes three parameters: two shape parameters, $\alpha > 0$ and $\beta > 0$, and a scale parameter, $\lambda > 0$. The PDF is given as follows:

$$g(t) = \frac{\lambda\beta}{\Gamma(\alpha)}(\lambda t)^{\alpha\beta-1}\,\mathrm{e}^{(-\lambda t)^{\beta}}, \quad t > 0 \tag{2}$$

The four-parameter mixture generalized gamma (MGG) distribution [7] was proposed by combining the generalized gamma distribution and its length-biased version. The MGG distribution consists of several sub-models such as generalized gamma, exponential, length-biased exponential and length-biased generalized gamma distributions. Nonetheless, the parameter estimation encounters issues. The greater the number of parameters, the more difficult and complicated the problem. Furthermore, Phaphan's study [8] found that the estimate of the weighted parameter p of the MGG distribution was outside the interval $[0, 1]$.

Afterwards, utilizing a new weighted parameter, Abdullahi and Phaphan [9] created a new MGG distribution by reducing the number of parameters to three. It was also a hybrid of the generalized gamma distribution and its length-biased form. The distribution was called a weighted mixture generalized gamma (WMGG) distribution. According to the results of [9], the WMGG distribution outperformed other candidate distributions in terms of goodness of fit. This indicates that the WMGG distribution is flexible and suitable for survival analysis as an alternative distribution. Abdullahi and Phaphan [9] also proposed the maximum likelihood estimators (MLEs) of WMGG distribution via quasi-Newton method. However, there were still issues with parameter estimation. Only the parameter λ showed good performance for the quasi-Newton method.

This motivated the prime objective of this research. Accordingly, the parameter estimation of the WMGG distribution was further developed. Since the concept of maximum likelihood is both straightforward and flexible, it has become a dominant method of statistical inference [10],[11],[12]. This article aims to propose the maximum likelihood estimation of the WMGG distribution employing three iterative approaches: EM algorithm, simulated annealing, and quasi-Newton for estimating the parameters λ, β, and α. A performance comparison of the three methods is evaluated via Monte Carlo simulation. In each scenario, the simulation is run 10,000 times. The bias and MSE are used as evaluation criteria.

The rest of the article is structured as follows. A description of the WMGG distribution is given in section 2. Three methods of maximum likelihood estimation are described in section 3. The simulation study procedure is explained in section 4. The numerical results are reported in section 5. Conclusion and discussion are finally presented in section 6.

2 Weighted Mixture Generalized Gamma Distribution

Abdullahi and Phaphan [9] introduced the weighted mixture generalized gamma (WMGG) distribution, which is the continuous probability distribution with three parameters, λ, β, and α. Suppose T is a random variable with the WMGG distribution. The PDF is denoted as

$$f(t) = \frac{\lambda}{\lambda+1}g(t) + \frac{1}{\lambda+1}g_L(t), \quad t > 0, \tag{3}$$

where

$$g(t) = \frac{\lambda\beta}{\Gamma(\alpha)}(\lambda t)^{\alpha\beta-1}\,\mathrm{e}^{-(\lambda t)^{\beta}}, \quad t > 0, \tag{4}$$

$$g_L(t) = \frac{\lambda\beta}{\Gamma\left(\alpha+\frac{1}{\beta}\right)}(\lambda t)^{\alpha\beta}\,\mathrm{e}^{(-\lambda t)^{\beta}}, \quad t > 0 \tag{5}$$

By inserting (4) and (5) into (3), hence the PDF of the WMGG distribution can be defined as

$$f(t) = \frac{\lambda\beta}{(\lambda+1)}(\lambda t)^{\alpha\beta}\,\mathrm{e}^{(-\lambda t)^{\beta}}\left(\frac{1}{t\Gamma(\alpha)} + \frac{1}{\Gamma\left(\alpha+\frac{1}{\beta}\right)}\right), \tag{6}$$

where $t > 0$, $\lambda > 0$, $\beta > 0$, and $\alpha > 0$. The cumulative distribution function or CDF of the WMGG distribution is expressed as

$$\begin{aligned} F(x) =& 1 - \left(\frac{\lambda}{\lambda+1}\right)\left[\frac{\Gamma\left(\alpha,(\lambda x)^{\beta}\right)}{\Gamma(\alpha)}\right] \\ &- \left(\frac{1}{\lambda+1}\right)\left[\frac{\Gamma\left(\alpha+\frac{1}{\beta},(\lambda x)^{\beta}\right)}{\Gamma\left(\alpha+\frac{1}{\beta}\right)}\right], \end{aligned} \tag{7}$$

here $\lambda > 0$, $\beta > 0$, and $\alpha > 0$.

In this article, the quantile function of the WMGG distribution, the further distributional function, is proposed as follows:

$$\begin{aligned} t =& \frac{1}{\lambda+1}\left[\gamma^{-1}\left(\alpha, u\Gamma(\alpha)\right)\right]^{\frac{1}{\beta}} \\ &+ \frac{1}{(\lambda+1)\lambda}\left[\gamma^{-1}\left(\alpha+\frac{1}{\beta}, u\Gamma\left(\alpha+\frac{1}{\beta}\right)\right)\right]^{\frac{1}{\beta}}, \end{aligned} \tag{8}$$

where γ^{-1} is the inverse incomplete gamma function, and u represents a uniform distribution.

3 Parameter Estimation Methods

The maximum likelihood estimation of the WMGG distribution's parameters was developed in this section. The MLE of the EM algorithm, simulated annealing, and quasi-Newton method based on a complete sample $t_1, t_2, ..., t_n$ were described, respectively.

3.1 MLE using EM Algorithm

An expectation-maximization (EM) algorithm is an iterative technique for determining maximum likelihood estimates of statistical model parameters. It is also used to find maximum a posteriori (MAP) estimate in Bayesian inference. The EM algorithm is commonly used for the following two reasons. Firstly, there are missing values due to limitations in the data collection process. Secondly, it is challenging to maximize the likelihood function under consideration. The EM algorithm consists of two steps. The expectation (E) step constructs the expectation of the log-likelihood function appraised by the current parameter estimates. The maximization (M) step calculates parameters achieved by maximizing the log-likelihood function's expectation from the previous step. The EM algorithm's procedures are given as follows.

The Procedure of E-Step

Step 1: Find the log-likelihood function of the random variable $T \sim WMGG(\lambda, \beta, \alpha)$ or $\ell(t)$. The function is given by

$$\ell(t) = \sum_{i=1}^{n} \ln \left\{ \frac{\lambda}{\lambda+1} g(t) + \frac{t\lambda\Gamma(\alpha)}{(\lambda+1)\,\Gamma\left(\alpha+\frac{1}{\beta}\right)} g(t) \right\} \quad (9)$$

Step 2: Compute the complete log-likelihood function by giving missing values κ_i into (9). Let $Y = (T; K)$ be the complete random variable such that $y_1, y_2, ..., y_n$ are $y_i = (t_i, \kappa_i)$, where κ_i equals 0 or 1 and $i = 1, 2, ..., n$. Consequently, the resulting function can be expressed as

$$\begin{aligned} \ell\left(\Theta | y_1, y_2, \cdots, y_n\right) = & \sum_{i=1}^{n} \kappa_i \ln \left[\frac{\lambda}{\lambda+1} g(t) \right] \\ & + \sum_{i=1}^{n} (1-\kappa_i) \\ & \times \ln \left[\frac{t}{\Gamma\left(\alpha+\frac{1}{\beta}\right)} \frac{\lambda\Gamma(\alpha)}{\lambda+1} g(t) \right] \end{aligned} \quad (10)$$

Substituting (4) into (10), the complete log-likelihood function, $\ell\left(\Theta | y_1, y_2, \cdots, y_n\right)$, is obtained as follows.

$$\begin{aligned} & \sum_{i=1}^{n} t_i - n\ln(\lambda+1) - n(\bar{\kappa}-1)\ln\Gamma\left(\alpha+\frac{1}{\beta}\right) \\ & + (\alpha\beta+1)\, n\ln(\lambda) - n\bar{\kappa}\ln\Gamma(\alpha) + n\ln(\beta) \\ & + (\alpha\beta-1)\sum_{i=1}^{n}\ln(t_i) - \lambda^{\beta}\sum_{i=1}^{n} t_i - \sum_{i=1}^{n}\kappa_i ln(t_i) \end{aligned} \quad (11)$$

Step 3: Formulate the new equation by deleting constant expressions and determining $\bar{\kappa} = \frac{1}{n}\sum_{i=1}^{n}\kappa_i$, the resulting complete log-likelihood function is given as

$$\begin{aligned} & - n\ln(\lambda+1) - n(\bar{\kappa}-1)\ln\Gamma\left(\alpha+\frac{1}{\beta}\right) \\ & + (\alpha\beta+1)\, n\ln(\lambda) - n\bar{\kappa}\ln\Gamma(\alpha) + n\ln(\beta) \\ & + (\alpha\beta-1)\sum_{i=1}^{n}\ln(t_i) - \lambda^{\beta}\sum_{i=1}^{n} t_i \end{aligned} \quad (12)$$

Step 4: Compute parameter estimates of λ, β, and α. Their MLEs were gained by maximizing $\ell\left(\Theta | y_1, y_2, \cdots, y_n\right)$ in (12) via the nlminb() function in R [13]. Note that the pseudo-log-likelihood function was obtained at the E-step by replacing missing values with their expectations. Therefore, the pseudo-log-likelihood function at the k^{th} state is given as

$$\begin{aligned} & - n\ln(\lambda+1) - n(a^{(k)}-1)\ln\Gamma\left(\alpha+\frac{1}{\beta}\right) \\ & + (\alpha\beta+1)\, n\ln(\lambda) - na^{(k)}\ln\Gamma(\alpha) + n\ln(\beta) \\ & + (\alpha\beta-1)\sum_{i=1}^{n}\ln(t_i) - \lambda^{\beta}\sum_{i=1}^{n} t_i, \end{aligned} \quad (13)$$

where $a^{(k)} = \frac{1}{n}\sum_{i=1}^{n} a_i^{(k)}$, and $a_i^{(k)}$ is given by

$$a_i^{(k)} = \frac{\frac{\lambda}{\lambda+1} g(t; \alpha^{(k)}, \beta^{(k)}, \lambda^{(k)})}{\frac{\lambda}{\lambda+1} g(t; \alpha^{(k)}, \beta^{(k)}, \lambda^{(k)}) + \frac{1}{\lambda+1} g_L(t; \alpha^{(k)}, \beta^{(k)}, \lambda^{(k)})} \quad (14)$$

The Procedure of M-Step

The M-step procedure involves continuously repeating the round in order to increase the number of function expressions. In each round, if the $a^{(k)}$ values change, the parameter estimates of λ, β, and α also change at the k^{th} round ($\lambda^{(k+1)}, \beta^{(k+1)}$, and $\alpha^{(k+1)}$). The process occurs continuously until the parameter estimates unchanged. Accordingly, the MLEs of λ, β, and α via an EM algorithm were obtained at that round by maximizing (17). The initial values used in the EM algorithm were sugested by [8]: $\lambda^{(0)}, \beta^{(0)}$, and $\alpha^{(0)}$ are

$$\begin{aligned} \lambda^{(0)} &= \frac{\bar{t}}{\sigma^2}, \\ \beta^{(0)} &= \frac{\bar{t}}{\alpha^0}, \\ \text{and } \alpha^{(0)} &= \frac{0.5}{log(\bar{t}) - mean\left(log(t)\right)}, \end{aligned} \quad (15)$$

where $\bar{t} = \frac{\sum_{i=1}^{n} t_i}{n}$, and $\sigma^2 = \sum_{i=1}^{n} \frac{(t_i-\bar{t})^2}{n}$.

Step 1: Generate a random sample $t_1, t_2, \cdots, t_n$ following the WMGG distribution.

Step 2: Set $k = 0$, compute $\lambda^{(0)}, \beta^{(0)}$, and $\alpha^{(0)}$ as given in (15).

Step 3: Calculate $a^{(k)} = \frac{1}{n}\sum_{i=1}^{n} a_i^{(k)}$ for $i = 1, 2, \cdots, n$, where $a_i^{(k)}$ was given in (16). For instance, if $k = 0$, we get

$$a_i^{(0)} = \frac{\frac{\lambda}{\lambda+1} g(t; \alpha^{(0)}, \beta^{(0)}, \lambda^{(0)})}{\frac{\lambda}{\lambda+1} g(t; \alpha^{(0)}, \beta^{(0)}, \lambda^{(0)}) + \frac{1}{\lambda+1} g_L(t; \alpha^{(0)}, \beta^{(0)}, \lambda^{(0)})}. \quad (16)$$

Step 4: Find $\lambda^{(k+1)}, \beta^{(k+1)}$, and $\alpha^{(k+1)}$ by maximizing (17). For $k=0$, we obtain

$$\begin{aligned} &- n\ln(\lambda+1) - n(a^{(0)}-1)\ln\Gamma\left(\alpha+\frac{1}{\beta}\right) \\ &+ (\alpha\beta+1)\, n\ln(\lambda) - na^{(0)}\ln\Gamma(\alpha) + n\ln(\beta) \\ &+ (\alpha\beta-1)\sum_{i=1}^{n}\ln(t_i) - \lambda^{\beta}\sum_{i=1}^{n} t_i, \end{aligned} \quad (17)$$

Step 5: Check the condition, if $\lambda^{(k+1)} = \lambda^{(k)}, \beta^{(k+1)} = \beta^{(k)}$, and $\alpha^{(k+1)} = \alpha^{(k)}$, then stop; otherwise update $k = k+1$ and go back to Step 3 and Step 4.

3.2 MLE using Simulated Annealing Algorithm

Simulated annealing is a probabilistic method for approximating a function. In particular, it is a metaheuristic to approximate certain functions for an optimization problem in a large search space. Simulated annealing algorithm procedure can be described as the following steps:

Step 1: Set the initial value $x^{(k=0)}$, T, n, and the accuracy ε.

Step 2: Pick a random point $x^{(k+1)}$ around point $x^{(k)}$.

Step 3: If $\Delta E < 0$, then $x^{(k+1)}$ is accepted as the current state, where $\Delta E = f(x^{(k+1)}) - f(x^{(k)})$, and $f(x)$ is the objective function; else if $\alpha \leq \exp(\frac{-\Delta E}{KT})$, $x^{(k+1)}$ is accepted, where K is known as the Boltzmann constant, T is the temperature, and α is a random number from the uniform distribution in case $\alpha \sim U(0,1)$; otherwise, go back to Step 2.

Step 4: If $|x^{(k+1)} - x^{(k)}| < \varepsilon$ and T is small enough, then the repetition is stopped; otherwise, update $k = k+1$ and go back to Step 2.

The MLEs using the simulated annealing in this work were obtained via the optim() function, the optimization function in R.

3.3 MLE using Quasi-Newton Method

In this section, the result of Abdullahi and Phaphan [9] is concisely described. Let $t_1, \cdots, t_n$ be a random sample of size n taken from the WMGG distribution, then the likelihood function of the WMGG distribution is given by

$$\begin{aligned} &L(t) \\ &= \prod_{i=1}^{n}\left\{\frac{\lambda\beta}{(\lambda+1)}(\lambda t)^{\alpha\beta}\mathrm{e}^{(-\lambda t)^{\beta}}\left(\frac{1}{t\Gamma(\alpha)}+\frac{1}{\Gamma\left(\alpha+\frac{1}{\beta}\right)}\right)\right\} \end{aligned} \quad (18)$$

The log-likelihood function is defined as

$$\begin{aligned} \ell(t) &= n\ln(\lambda\beta) - n\ln(\lambda+1) + \alpha\beta\sum_{i=1}^{\infty}\ln(\lambda t_i) \\ &- \lambda^{\beta}\sum_{i=1}^{\infty} t_i^{\beta} + \sum_{i=1}^{\infty}\ln\left(\frac{1}{t\Gamma(\alpha)}+\frac{1}{\Gamma\left(\alpha+\frac{1}{\beta}\right)}\right) \end{aligned} \quad (19)$$

The derivative equations corresponding to the parameters, α, λ and β are presented as follows.

$$\begin{aligned} \frac{\partial\ell(t)}{\partial\alpha} &= \beta\sum_{i=1}^{\infty}\ln(\lambda t_i) \\ &- \sum_{i=1}^{\infty}\left(\frac{t_i\Gamma(\alpha)\Psi\left(\alpha+\frac{1}{\beta}\right)+\Gamma\left(\alpha+\frac{1}{\beta}\right)\Psi(\alpha)}{t_i\Gamma(\alpha)+\Gamma\left(\alpha+\frac{1}{\beta}\right)}\right), \end{aligned} \quad (20)$$

where

$$\begin{aligned} \frac{\partial ln(\Gamma(\alpha))}{\partial\alpha} &= \Psi(\alpha), \\ \frac{\partial\Gamma(\alpha)}{\partial\alpha} &= \Psi(\alpha)\Gamma(\alpha) \end{aligned} \quad (21)$$

$$\frac{\partial\ell(t)}{\partial\lambda} = \frac{n}{\lambda} - \frac{n}{\lambda+1} + \frac{n\alpha\beta}{\lambda} - \beta\lambda^{\beta-1}\sum_{i=1}^{\infty} t_i^{\beta}, \quad (22)$$

$$\begin{aligned} \frac{\partial\ell(t)}{\partial\beta} &= \frac{n}{\beta} + \alpha\sum_{i=1}^{\infty}\ln(\lambda t_i) - \sum_{i=1}^{\infty}(\lambda t_i)^{\beta}\ln(\lambda t_i) \\ &+ \sum_{i=1}^{\infty}\frac{t_i\Gamma(\alpha)\Psi\left(\alpha+\frac{1}{\beta}\right)}{t_i\beta^2\Gamma(\alpha)+\beta^2\Gamma\left(\alpha+\frac{1}{\beta}\right)} \end{aligned} \quad (23)$$

Since these equations are nonlinear, iterative methods can be used to solve them numerically rather than analytically. The MLEs using the quasi-Newton method of the three parameters can be obtained through the nlminb() function in R.

4 Simulation Study

The Monte Carlo simulation was established in this section to evaluate the performance of the proposed estimators using R [13]. The sample sizes varied from small to large, and various true parameter values were considered. All 72 situations were created from combinations of sample size $(n) = 10, 30, 50, 100$, parameter values $\lambda = 0.7, 1.4$, $\beta = 1$, and $\alpha = 3, 5, 7$, and three parameter estimation methods.

The random number for the WMGG distribution was generated using the acceptance-rejection method. This technique is an effective algorithm for generating mixture distributions. The basic idea of this technique is to find an alternative probability distribution of $T \sim WMGG(\lambda,\beta,\alpha)$, with the PDF, $f(t)$. Let G denote the alternative distribution with the PDF, $g(x)$, in case where the function $g(x)$ approaches to $f(x)$. Precisely, the ratio $f(x)/g(x)$ is treated to be bounded by a positive constant c, where $\sup_x f(x)/g(x) \leq c$. The constant c should be as close to 1 as possible. For our alternative distribution, the log-normal, $X \sim LN(\mu,\sigma^2)$, is selected since it is one of the special cases of the WMGG distribution. The main steps for generating the WMGG random numbers can be explained as follows:

Step 1: Create a random number X that has a log-normal distribution with the following PDF:

$$g(x) = \frac{1}{x\sigma\sqrt{2\pi}}\exp\left(-\frac{1}{2\sigma^2}\left(lnx-\mu\right)^2\right), \quad t>0,$$

where μ and σ are mean and standard deviation, respectively. The dlnorm() function was used to create the log-normal density, $g(x)$.
Step 2: Construct a random number U that has a uniform distribution, $U \sim U(0,1)$. The random variables U and X are independent. The runif(n, 0, 1) function was implemented to generate U.
Step 3: If

$$U(0,1) \leq \frac{f(x)}{cg(x)},$$

then set $T = X$ ("accept"); otherwise go back to Step1 ("reject").

Tables 1-12 report the parameter estimates, along with their bias and mean squared errors (MSE). In each table, six distribution models were created with different parameter values (λ, β, and α). Simulations were run 10,000 times in each model. The average values of parameter estimates, bias values, and MSEs were computed sequentially. Let $\boldsymbol{\theta}$ be the vector of the parameters λ, β, and α or $\boldsymbol{\theta}' = [\lambda, \beta, \alpha]'$.

The $\hat{\boldsymbol{\theta}}$ indicates the MLEs using EM algorithm. The $\tilde{\boldsymbol{\theta}}$ represents the MLEs using Simulated Annealing algorithm. The $\breve{\boldsymbol{\theta}}$ displays the MLEs using Quasi-Newton method. Bar charts for bias comparison were built and shown in Figures 1-3 for simpler consideration. The EM algorithm, simulated annealing, and quasi-Newton method estimator performances are shown, respectively, by the red, blue, and green bars.

5 Numerical Results

According to Tables 1-12 and Figures 1-3, it shows that as the sample size increases, the bias and MSE decrease, and the estimates get closer to the true value of the parameters. Interestingly, the EM algorithm outperforms other candidate algorithms for all parameters in terms of bias consideration. The EM algorithm's parameter estimates are reasonably close to the true values.

For parameter λ, the EM algorithm provides over-estimates when sample size is small; while, it offers under-estimates when sample size is large. The simulated annealing algorithm gives under-estimates for all scenarios and it is better than the quasi-Newton method. In conversely, the quasi-Newton method produces over-estimates for all scenarios, with estimates that are far from the true values.

For parameter β, the EM algorithm and quasi-Newton method offer under-estimates for all scenarios. Moreover, the EM algorithm is slightly better than the quasi-Newton method. In contrast, the simulated annealing algorithm provides over-estimates for all models, and its estimates are far from the actual values.

For parameter α, the EM algorithm gives over-estimates for all scenarios. When the sample gets bigger, the estimates become more accurate. Surprisingly, the simulated annealing and quasi-Newton methods perform identically.

When the sample size is moderate or large, the MSE criterion behaves similarly to the bias criterion. Nonetheless, when the sample size is small, the performance of each method and parameter differs.

For parameter λ, the simulated annealing algorithm outperforms other methods. However, the simulated annealing algorithm is slightly better than the EM algorithm. The quasi-Newton method yields the greatest MSE.

For parameter β, on the other hand, the simulated annealing algorithm performs poorly. The EM algorithm is the most superior method, followed by the quasi-Newton method.

For parameter α, the EM algorithm outperforms other methods for all scenarios. Like the bias criterion, the performance of the simulated annealing and quasi-Newton methods is identical.

6 Conclusion and Discussion

The WMGG distribution introduced by [9] has useful properties. However, the parameter estimation runs into issues. Therefore, the parameter estimation should be further improved. Maximum likelihood estimation was selected because of straightforward and flexible properties. The two iterative methods, EM algorithm and simulated annealing, were proposed for estimating the WMGG distribution's parameters in this article. The two methods were compared to the quasi-Newton method suggested by [9] via Monte Carlo simulation. Evaluation criteria were MSE and bias. Regardless of parameter estimation, the quantile function of the WMGG distribution was presented in this paper.

The numerical results indicate that, overall, the EM algorithm outperformed the candidate methods, followed by the simulated annealing algorithm, and the least effective method is the quasi-Newton one. It's interesting to note that the estimators obtained by the proposed methods outperform the estimators from the previous study [9]. Our contribution provides a significant result. The EM algorithm, especially, provides parameter estimates that are close to the true parameter values. Consequently, one can apply this method to other complicated functions.

The Bayesian methods such as [14]-[15] or other techniques should be considered for parameter estimation in future work.

Acknowledgements

The authors would like to acknowledge the referees for their insightful comments on the manuscript. This research was funded by King Mongkut's University of Technology North Bangkok, Contract no. KMUTNB-65-BASIC-16.

Table 1. Maximum Likelihood estimates and their bias and MSE from the EM algorithm for $n = 10$

λ	β	α	$\hat{\lambda}$	$\hat{\beta}$	$\hat{\alpha}$	B($\hat{\lambda}$)	B($\hat{\beta}$)	B($\hat{\alpha}$)	MSE($\hat{\lambda}$)	MSE($\hat{\beta}$)	MSE($\hat{\alpha}$)
0.7	1	3	0.73	0.52	4.26	0.03	-0.48	1.26	0.48	1.19	10.33
0.7	1	5	0.76	0.4	7.14	0.06	-0.6	2.14	0.44	1.11	25.88
0.7	1	7	0.8	0.28	10.09	0.1	-0.72	3.09	0.43	1.05	53.14
1.4	1	3	1.74	0.04	4.19	0.34	-0.96	1.19	1.43	0.97	8.88
1.4	1	5	1.76	0.02	7.06	0.36	-0.98	2.06	1.54	0.98	25.55
1.4	1	7	1.77	0.01	9.94	0.37	-0.99	2.94	1.4	0.99	48.09

Table 2. Maximum Likelihood estimates and their bias and MSE from the simulated annealing for $n = 10$

λ	β	α	$\tilde{\lambda}$	$\tilde{\beta}$	$\tilde{\alpha}$	B($\tilde{\lambda}$)	B($\tilde{\beta}$)	B($\tilde{\alpha}$)	MSE($\tilde{\lambda}$)	MSE($\tilde{\beta}$)	MSE($\tilde{\alpha}$)
0.7	1	3	0.42	27.52	28.43	-0.28	26.52	25.43	0.25	902.03	907.96
0.7	1	5	0.19	33.61	34.07	-0.51	32.61	29.07	0.28	1368.93	1221.53
0.7	1	7	0.12	35.97	39.11	-0.58	34.97	32.11	0.34	1553.78	1495.22
1.4	1	3	0.91	26.32	28.5	-0.49	25.32	25.5	4.26	800.68	909.66
1.4	1	5	0.44	28.96	35.37	-0.96	27.96	30.37	1.35	988.4	1304.1
1.4	1	7	0.27	32.08	41.19	-1.13	31.08	34.19	1.37	1218.44	1642.62

Table 3. Maximum Likelihood estimates and their bias and MSE from the quasi-Newton method for $n = 10$

λ	β	α	$\breve{\lambda}$	$\breve{\beta}$	$\breve{\alpha}$	B($\breve{\lambda}$)	B($\breve{\beta}$)	B($\breve{\alpha}$)	MSE($\breve{\lambda}$)	MSE($\breve{\beta}$)	MSE($\breve{\alpha}$)
0.7	1	3	4.86	-2.32	28.43	4.16	-3.32	25.43	55.42	175.62	907.96
0.7	1	5	4.95	-2.16	34.07	4.25	-3.16	29.07	54.32	173.37	1221.53
0.7	1	7	4.92	-1.83	39.11	4.22	-2.83	32.11	52.5	155.23	1495.22
1.4	1	3	4.48	-3.23	28.5	3.08	-4.23	25.5	50.42	202.53	909.66
1.4	1	5	4.84	-2.14	35.37	3.44	-3.14	30.37	51.98	172.66	1304.1
1.4	1	7	4.98	-2.14	41.19	3.58	-3.14	34.19	52.59	179.33	1642.62

Table 4. Maximum Likelihood estimates and their bias and MSE from the EM algorithm for $n = 30$.

λ	β	α	$\hat{\lambda}$	$\hat{\beta}$	$\hat{\alpha}$	B($\hat{\lambda}$)	B($\hat{\beta}$)	B($\hat{\alpha}$)	MSE($\hat{\lambda}$)	MSE($\hat{\beta}$)	MSE($\hat{\alpha}$)
0.7	1	3	0.55	0.37	3.22	-0.15	-0.63	0.22	0.18	0.9	1.3
0.7	1	5	0.61	0.18	5.52	-0.09	-0.82	0.52	0.13	0.93	3.42
0.7	1	7	0.66	0.11	7.8	-0.04	-0.89	0.8	0.11	0.95	6.27
1.4	1	3	1.41	0.01	3.26	0.01	-0.99	0.26	0.22	0.99	1.07
1.4	1	5	1.46	0	5.54	0.06	-1	0.54	0.21	0.99	2.88
1.4	1	7	1.47	0	7.75	0.07	-1	0.75	0.2	0.99	5.43

Table 5. Maximum Likelihood estimates and their bias and MSE from the simulated annealing for $n = 30$.

λ	β	α	$\tilde{\lambda}$	$\tilde{\beta}$	$\tilde{\alpha}$	B($\tilde{\lambda}$)	B($\tilde{\beta}$)	B($\tilde{\alpha}$)	MSE($\tilde{\lambda}$)	MSE($\tilde{\beta}$)	MSE($\tilde{\alpha}$)
0.7	1	3	0.42	19.6	28.52	-0.28	18.6	25.52	0.43	383.24	740.94
0.7	1	5	0.19	22.61	32.95	-0.51	21.61	27.95	0.28	502.95	867.54
0.7	1	7	0.13	20.79	39.77	-0.57	19.79	32.77	0.33	415.62	1161.62
1.4	1	3	0.84	20.76	26.94	-0.56	19.76	23.94	2.07	437.8	667.78
1.4	1	5	0.44	20.94	34.37	-0.96	19.94	29.37	1.25	453.22	963.21
1.4	1	7	0.27	20.94	39.67	-1.13	19.94	32.67	1.32	436.52	1160.16

Table 6. Maximum Likelihood estimates and their bias and MSE from the quasi-Newton method for $n = 30$.

λ	β	α	$\breve{\lambda}$	$\breve{\beta}$	$\breve{\alpha}$	B($\breve{\lambda}$)	B($\breve{\beta}$)	B($\breve{\alpha}$)	MSE($\breve{\lambda}$)	MSE($\breve{\beta}$)	MSE($\breve{\alpha}$)
0.7	1	3	4.43	-2.28	28.52	3.73	-3.28	25.52	50.18	165.7	740.94
0.7	1	5	4.73	-1.57	32.95	4.03	-2.57	27.95	50.01	143.74	867.54
0.7	1	7	4.84	-1.71	39.77	4.14	-2.71	32.77	50.87	150.78	1161.62
1.4	1	3	3.48	-4.62	26.94	2.08	-5.62	23.94	35.81	201.08	667.78
1.4	1	5	4.37	-2.04	34.37	2.97	-3.04	29.37	45.77	155.15	963.21
1.4	1	7	4.65	-1.52	39.67	3.25	-2.52	32.67	47.48	149.64	1160.16

Table 7. Maximum Likelihood estimates and their bias and MSE from the EM algorithm for $n = 50$.

λ	β	α	$\hat{\lambda}$	$\hat{\beta}$	$\hat{\alpha}$	B($\hat{\lambda}$)	B($\hat{\beta}$)	B($\hat{\alpha}$)	MSE($\hat{\lambda}$)	MSE($\hat{\beta}$)	MSE($\hat{\alpha}$)
0.7	1	3	0.51	0.28	3.03	-0.19	-0.72	0.03	0.15	0.85	0.69
0.7	1	5	0.61	0.11	5.28	-0.09	-0.89	0.28	0.09	0.93	1.69
0.7	1	7	0.65	0.05	7.45	-0.05	-0.95	0.45	0.06	0.96	3.06
1.4	1	3	1.36	0	3.13	-0.04	-1	0.13	0.12	0.99	0.53
1.4	1	5	1.39	0	5.27	-0.01	-1	0.27	0.11	0.99	1.4
1.4	1	7	1.41	0	7.42	0.01	-1	0.42	0.11	0.99	2.7

Table 8. Maximum Likelihood estimates and their bias and MSE from the simulated annealing for $n = 50$.

λ	β	α	$\tilde{\lambda}$	$\tilde{\beta}$	$\tilde{\alpha}$	B($\tilde{\lambda}$)	B($\tilde{\beta}$)	B($\tilde{\alpha}$)	MSE($\tilde{\lambda}$)	MSE($\tilde{\beta}$)	MSE($\tilde{\alpha}$)
0.7	1	3	0.42	17.54	29.57	-0.28	16.54	26.57	1.11	304.02	780.01
0.7	1	5	0.2	19.16	34.75	-0.5	18.16	29.75	0.3	350.96	959.4
0.7	1	7	0.13	18.32	40.61	-0.57	17.32	33.61	0.33	317.77	1191.11
1.4	1	3	0.89	19.37	27.07	-0.51	18.37	24.07	13.85	380.04	668.44
1.4	1	5	0.44	18.31	35.23	-0.96	17.31	30.23	1.35	333.93	992.72
1.4	1	7	0.28	17.63	41.19	-1.12	16.63	34.19	1.31	293.88	1232.7

Table 9. Maximum Likelihood estimates and their bias and MSE from the quasi-Newton method for $n = 50$.

λ	β	α	$\breve{\lambda}$	$\breve{\beta}$	$\breve{\alpha}$	B($\breve{\lambda}$)	B($\breve{\beta}$)	B($\breve{\alpha}$)	MSE($\breve{\lambda}$)	MSE($\breve{\beta}$)	MSE($\breve{\alpha}$)
0.7	1	3	4.2	-2.15	29.57	3.5	-3.15	26.57	46.53	158.11	780.01
0.7	1	5	4.57	-1.52	34.75	3.87	-2.52	29.75	48.53	145.49	959.4
0.7	1	7	4.75	-1.24	40.61	4.05	-2.24	33.61	50.33	134.43	1191.11
1.4	1	3	3.02	-5.37	27.07	1.62	-6.37	24.07	28.56	202.13	668.44
1.4	1	5	4.12	-2.32	35.23	2.72	-3.32	30.23	44.42	154.05	992.72
1.4	1	7	4.49	-1.69	41.19	3.09	-2.69	34.19	46.84	147.97	1232.7

Table 10. Maximum Likelihood estimates and their bias and MSE from the EM algorithm for $n = 100$.

λ	β	α	$\hat{\lambda}$	$\hat{\beta}$	$\hat{\alpha}$	B($\hat{\lambda}$)	B($\hat{\beta}$)	B($\hat{\alpha}$)	MSE($\hat{\lambda}$)	MSE($\hat{\beta}$)	MSE($\hat{\alpha}$)
0.7	1	3	0.49	0.16	2.88	-0.21	-0.84	-0.12	0.13	0.88	0.41
0.7	1	5	0.61	0.04	5.1	-0.09	-0.96	0.1	0.05	0.96	0.76
0.7	1	7	0.65	0.01	7.24	-0.05	-0.99	0.24	0.03	0.98	1.36
1.4	1	3	1.32	0	3.02	-0.08	-1	0.02	0.06	0.99	0.24
1.4	1	5	1.36	0	5.12	-0.04	-1	0.12	0.05	0.99	0.6
1.4	1	7	1.38	0	7.19	-0.02	-1	0.19	0.05	0.99	1.15

Table 11. Maximum Likelihood estimates and their bias and MSE from the simulated annealing for $n = 100$.

λ	β	α	$\tilde{\lambda}$	$\tilde{\beta}$	$\tilde{\alpha}$	B($\tilde{\lambda}$)	B($\tilde{\beta}$)	B($\tilde{\alpha}$)	MSE($\tilde{\lambda}$)	MSE($\tilde{\beta}$)	MSE($\tilde{\alpha}$)
0.7	1	3	0.4	15.11	31.26	-0.3	14.11	28.26	0.41	218.11	855.3
0.7	1	5	0.2	15.79	36.71	-0.5	14.79	31.71	0.26	230.34	1054.37
0.7	1	7	0.14	16.36	40.53	-0.56	15.36	33.53	0.32	247.59	1170.6
1.4	1	3	0.88	18.26	26.72	-0.52	17.26	23.72	4.7	338.82	645.31
1.4	1	5	0.43	15.56	36.82	-0.97	14.56	31.82	1.14	227.7	1066.33
1.4	1	7	0.28	16.06	40.84	-1.12	15.06	33.84	1.27	240.56	1188.33

Table 12. Maximum Likelihood estimates and their bias and MSE from the quasi-Newton method for $n = 100$.

λ	β	α	$\breve{\lambda}$	$\breve{\beta}$	$\breve{\alpha}$	B($\breve{\lambda}$)	B($\breve{\beta}$)	B($\breve{\alpha}$)	MSE($\breve{\lambda}$)	MSE($\breve{\beta}$)	MSE($\breve{\alpha}$)
0.7	1	3	3.82	-3.01	31.26	3.12	-4.01	28.26	43.87	166.41	855.3
0.7	1	5	4.44	-1.23	36.71	3.74	-2.23	31.71	49.5	134.71	1054.37
0.7	1	7	4.66	-1.27	40.53	3.96	-2.27	33.53	51.31	134.63	1170.6
1.4	1	3	2.48	-7.39	26.72	1.08	-8.39	23.72	23.77	222.89	645.31
1.4	1	5	3.77	-2.61	36.82	2.37	-3.61	31.82	41.92	152.38	1066.33
1.4	1	7	4.21	-1.51	40.84	2.81	-2.51	33.84	45.23	143.91	1188.33

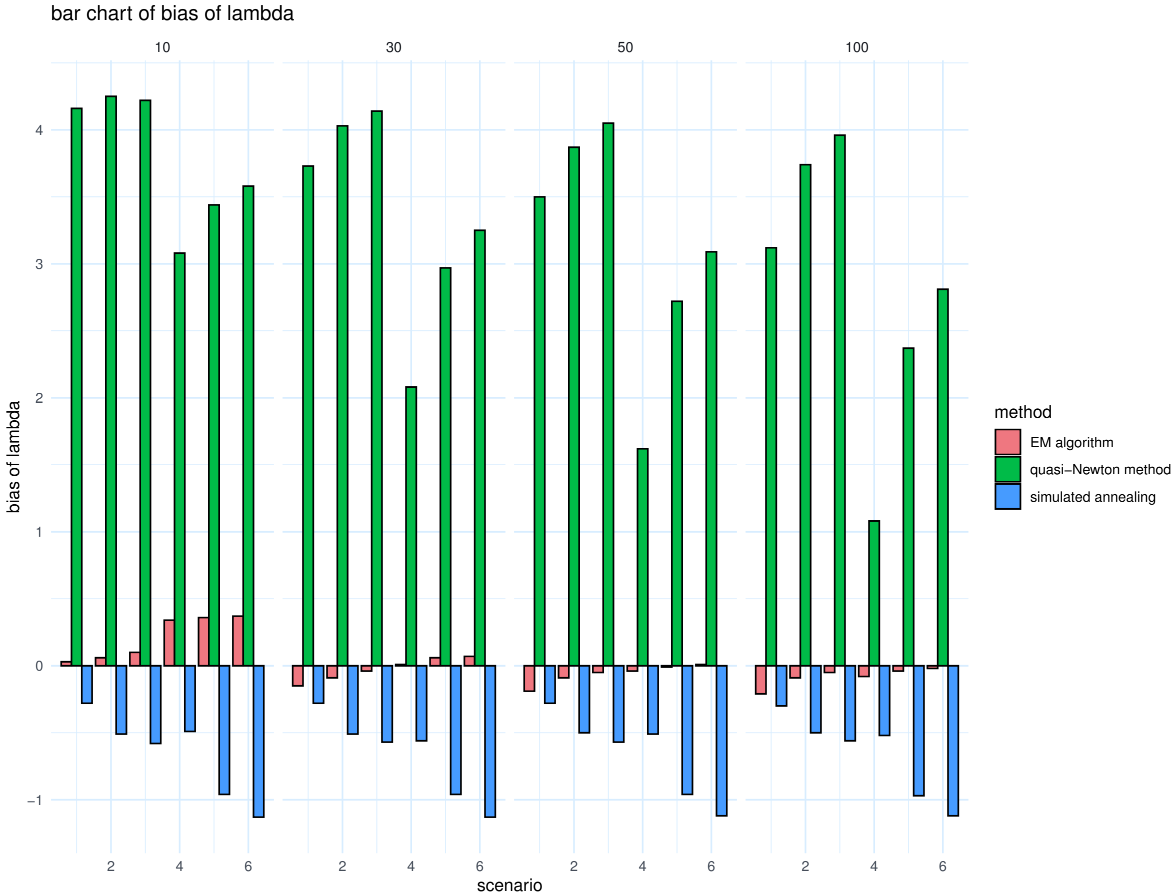

Figure 1. Bias bar chart for λ estimators

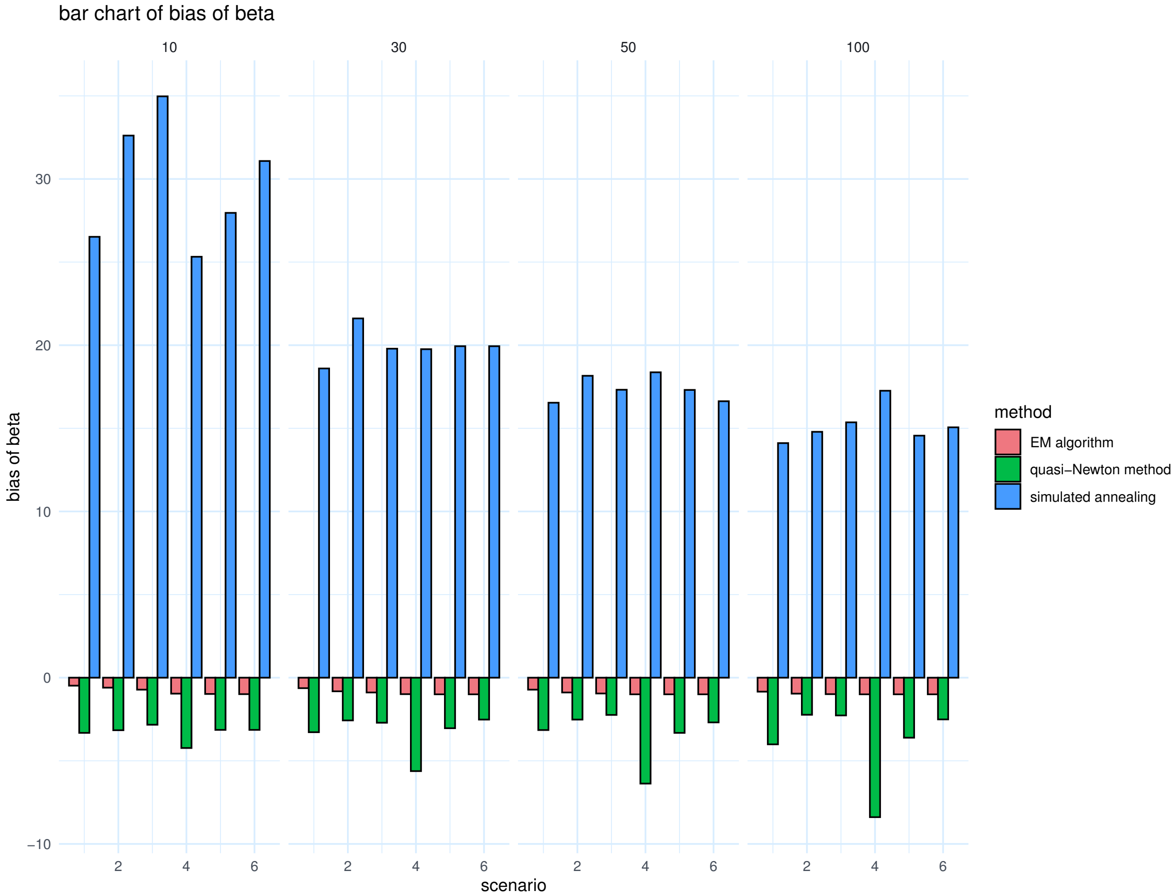

Figure 2. Bias bar chart for β estimators

Figure 3. Bias bar chart for α estimators

REFERENCES

[1] Erlang A. K., ”The theory of probabilities and telephone conversations,” Nyt Tidsskrift for Matematik B, vol. 20, No. 6, pp. 33-39, 1909.

[2] Stacy E.W., ”A generalization of the gamma distribution,” The Annals of Mathematical statistics, vol. 33, no. 3, pp. 1187-1192, 1962. http://www.jstor.org/stable/2237889.

[3] Favero C.A., Pesaran M.H., Sharma S., ”A duration model of irreversible oil investment: theory and empirical evidence,” Journal of Applied Econometrics, vol. 9, no. S1, pp. S95-S112, 1994. DOI: https://doi.org/10.1002/jae.3950090506

[4] Eckstein Z., Wolpin K.I., ”Duration to first job and the return to schooling: estimates from a search-matching model,” The Review of Economic Studies, vol. 62, no. 2, pp. 263-268, 1995. DOI: https://doi.org/10.2307/2297805

[5] Lancaster T., ”Econometric methods for the duration of unemployment,” Econometrica, Vol.47, no. 4, pp. 939-956, 1979. DOI: https://doi.org/10.2307/1914140

[6] Rayleigh O.M. F.R.S. L., ”XXXI. On the problem of random vibrations, and of random flights in one, two, or three dimensions,” The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, Vol.37, no. 220, pp. 321-347, 1919. DOI: 10.1080/14786440408635894

[7] Suksaengrakcharoen S., Bodhisuwan W., ”A new family of generalized gamma distribution and its application,” Journal of Mathematics and Statistics, vol. 10, no. 2, pp. 211-220, 2014. DOI: https://doi.org/10.3844/jmssp.2014.211.220

[8] Phaphan W., ”Estimating parameter for the mixture generalized gamma distribution,” The 10th International Conference on Computer Modeling and Simulation, Sydney, Australia, Jan. 2018, pp. 77-81. DOI: https://doi.org/10.1145/3177457.3177492

[9] Abdullahi I., Phaphan W., ”Some properties of the new mixture of generalized gamma distribution,” Lobachevskii Journal of Mathematics, vol.43, no. 9, pp. 2349–2359, 2022. DOI: https://doi.org/10.1134/S1995080222120022

[10] Hendry, D.F., Nielsen, B., ”Econometric Modeling: A Likelihood Approach,” PUP, 2007.

[11] Chambers, R.L., Steel, D.G., Wang, S., Welsh, A., ”Maximum Likelihood Estimation for Sample Surveys,” CRC Press, 2012.

[12] Ward, M.D., Ahlquist, J.S., ”Maximum Likelihood for Social Science : Strategies for Analysis,” CUP, 2018.

[13] R Core Team, ”R: A language and environment for statistical computing,” 2021, http://Www. R-Project. Org

[14] Chaiprasithikul, D., Duangsaphon, M., ”Bayesian Inference for the Discrete Weibull Regression Model with Type-I Right Censored Data,” Thailand Statistician, vol.20, no. 4, pp. 791-811, 2022.

[15] Chaiprasithikul, D., Duangsaphon, M., ”Bayesian Inference of Discrete Weibull Regression Model for Excess Zero Counts,” Science and Technology Asia, vol.27, no. 4, pp. 152-174, 2022.