



---

*Research article*

## The multi-parameter estimation of discrete distribution without closed-form solutions by the US algorithm

Yuanhang Ouyang<sup>1</sup>, Ruyun Yan<sup>1</sup> and Jianhua Shi<sup>1,2,\*</sup>

<sup>1</sup> School of Mathematics and Statistics, Minnan Normal University, Zhangzhou, China

<sup>2</sup> Fujian Key Laboratory of Granular Computing and Applications, Zhangzhou, China

\* **Correspondence:** Email: v0085@126.com.

**Abstract:** Strong and steady convergence characterizes the upper-crossing/solution (US) algorithm, which is an effective method for identifying roots of a complicated nonlinear equation  $h(\theta) = 0$ . Only the case where one parameter of a distribution function can be directly specified by another parameter is taken into account by the research that is currently available. However, whether this approach can be applied in multi-parameter scenarios where one parameter cannot be clearly represented by the other is an issue deserving of more investigation. In order to extend the applicability of the US algorithm, this article used the Type I discrete Weibull distribution with two parameters as an example. It then combined the US algorithm with the first-derivative lower bound (FLB) function method to estimate the complex situation where two parameters cannot be expressed as each other. Simulation studies and empirical analysis demonstrated that the US algorithm performs more accurately and steadily than the traditional Newton method.

**Keywords:** Upper-crossing/solution (US) algorithm; first-derivative lower bound (FLB) function method; maximum likelihood estimate; Type I discrete Weibull distribution

---

### 1. Introduction

Discrete lifetime data—such as the number of appliance failures of a particular brand within a given time frame, the total number of machine operations prior to a failure, the number of bullets fired by a weapon before the first malfunction, and the anticipated lifespan of humans (in years)—are frequently handled in reliability lifetime studies. For more classic examples, see Szymkowiak and Iwinska [1]. Data scientists typically employ discrete models as analysis tools, such as the Poisson distribution, negative binomial distribution, and geometric distribution, in order to more correctly define, analyze, and model these data. But in many situations, these discrete distribution functions are not the best options. For instance, seasonal or periodic data cannot be handled by the Poisson distribution, while

underdispersed data cannot be described by the negative binomial distribution. More suitable discrete lifetime distributions are required to explore many additional kinds of complex discrete lifespan data. Discretizing continuous random variables is a useful strategy that yields a discrete life model with characteristics that are comparable to the continuous model.

The essential concept of discretizing continuous random variables was first presented by Roy [2]. Specifically, let  $Y$  be a continuous random variable with a survival function denoted by  $S(y)$ . Define the random variable  $Z = [Y]$  as the maximum integer less than or equal to  $Y$ . The probability mass function (PMF)  $P(Z = z)$  of  $Z$  can be expressed as

$$P(Z = z) = S_Y(z) - S_Y(z + 1).$$

Many researchers have introduced various new models for discrete life distributions by the approach. For instance, the discrete normal distribution was first introduced by Roy [3]. Using the generic method of discretizing a continuous distribution, Krishna and Pundir [4] introduced the discrete Burr and Pareto distributions. In addition, Bracquemond and Gaudoin [5] provided an extensive overview of discrete distributions, such as the Weibull distribution, that are employed in reliability to describe discrete lifetimes of nonrepairable systems. It is well-known that the Weibull distribution has become the most commonly used distribution for analyzing continuous life data due to its ability to fit various types of data and relatively simple structure (Johnson et al. [6]). At least three cases exist for the corresponding discrete Weibull distribution: (a) the Type I discrete Weibull distribution, which maintains the form of the continuous survival function (SF), as introduced by Nakagawa and Osaki [7]; (b) the Type II discrete Weibull distribution, as suggested by Stein and Dattero [8]; and (c) the three-parameter discrete Weibull distribution, as introduced by Padgett and Spurrier [9]. The most popular of them is the Type I discrete Weibull distribution, whose features have been extensively researched by numerous academics. Englehardt and Li [10] employed the discrete Weibull distribution to analyze pathogen counts in treated water over time. Barbiero [11, 12] compared several parameter estimation methods of this distribution, and solved the minimum Chi-square and least squares estimation. Vila et al. [13] studied in detail the basic theoretical properties of the Type I discrete Weibull and analyzed the censored data. Yoo [14] extended the application of the discrete Weibull regression model to accommodate missing data. In addition, El-Morshedy et al. [15] conducted a detailed study on a new bivariate exponential discrete Weibull distribution.

The primary goal in this study is to enhance the current techniques for estimating the complex discrete probability distribution model. The probability distribution's score function typically lacks an explicit analytical solution, hence the Newton approach is usually used to estimate the numerical solution for parameter estimation. Nevertheless, the algorithm's low convergence and strong dependence on the initial value make it challenging to achieve the best estimation outcomes. Recently, Liu et al. [16] employed the majorize minimize (MM) algorithm to enhance the resolution of the maximum likelihood estimation for the simplex distribution. Li and Tian [17] introduced a novel root-finding method known as the upper-crossing/solution (US) algorithm. In contrast to conventional iterative algorithms (like Newton's algorithm), the US algorithm can lessen the influence of initial values and achieve a strong, stable convergence to the objective equation's real root at each iteration. The benefits of this technique have been illustrated through the use of a few classic models, such as the Weibull distribution, gamma distribution, zeta distribution, and generalized Poisson

distribution. Cai [18] has improved the maximum likelihood estimation of generalized gamma distribution parameters by combining the US algorithm with the second-derivative lower-bound function (SeLF) algorithm.

The essence of the US algorithm is to identify a  $U$ -function  $U(\theta|\theta^{(t)})$ , which simplifies the solution of the complicated nonlinear equation  $h(\theta) = 0$  to the solution of the equation  $U(\theta|\theta^{(t)}) = 0$  with an explicit solution. Li and Tian [17] presented a variety of approaches to discover the  $U$ -function, among which the first-derivative lower bound (FLB) function method requires only by using the first derivative of the objective function  $h(\theta)$ , thereby diminishing algorithmic complexity. In previous research on the US algorithm, it was generally used to solve the roots of univariate nonlinear equations or the maximum likelihood estimation problem of a multi-parameter probability distribution with an explicit partial score function. Specifically, for a probability distribution with two parameters  $(\alpha, \beta)$ , while solving for maximum likelihood estimation, the estimator of the parameter  $\alpha$  can be explicitly expressed by the other estimator of the parameter  $\beta(\alpha)$ . However, there is no further discussion provided in Li and Tian [17] about whether this approach can be applicable in more complex multi-parameter distributions, where an estimator of one parameter cannot be clearly represented by the other, and it is an issue deserving of more investigation.

The rest of the paper will proceed as follows. Section 2 provides a detailed introduction to the US algorithm and FLB function method. The application of the suggested approach to the Type I discrete Weibull distribution's maximum likelihood parameter estimation is covered in Section 3. Numerical simulation experiments will be conducted in Section 4 in order to evaluate the performance of the employed methods and compare them with alternative estimation approaches. Section 5 will demonstrate the applicability of the US algorithm through the analysis of two real data sets. Conclusions and discussions will be provided in Section 6.

## 2. The US algorithm

One of the most frequent issues in numerical computations is figuring out the zero point of a function or an equation's root. In classical statistics, the maximum likelihood estimate (MLE) of parameters and the calculation of maximum a posteriori probability in Bayesian statistics may typically be turned into the problem of solving the zero point of a nonlinear function  $h(\theta)$ . In summary, since  $h(\theta)$  is a nonlinear function of a single variable  $\theta$ , we must identify the unique root  $\theta^*$  such that

$$h(\theta) = 0, \quad \theta \in \Theta \subseteq \mathbb{R}. \quad (2.1)$$

The US algorithm is the most recent method for discovering roots. It has a similar procedure to the commonly used EM (expectation maximum) and MM (maximize minimize) algorithms [17]. There are two primary steps in this process: the upper-crossing step (U-step) and the solution step (S-step). The two primary advantages of this algorithm are as follows:

(a) It converges strongly and stably to the root  $\theta^*$  of the Eq (2.1) with each iteration, that is, for an iterative points set sequence  $\{\theta^{(t)}\}_{t=0}^{\infty}$ , there is

$$\theta^{(0)} < \theta^{(1)} < \dots < \theta^{(t)} < \dots \leq \theta^* \quad \text{or} \quad \theta^* \leq \dots < \theta^{(t)} < \dots < \theta^{(1)} < \theta^{(0)}.$$

(b) The Newton algorithm's sensitivity to the initial value is decreased.

Two new symbols,  $\overset{\text{sgn}(\alpha)}{\leq}$  and  $\overset{\text{sgn}(\alpha)}{\geq}$ , are introduced regarding the changing direction (CD) inequalities in order to simplify the explanation of the US algorithm. The specific definition is presented as follows: for two functions  $f_1(x)$  and  $f_2(x)$  on the same domain  $\mathbb{Q}$ ,

$$f_1(x) \overset{\text{sgn}(\alpha)}{\leq} f_2(x) \Leftrightarrow \begin{cases} f_1(x) \leq f_2(x), & \alpha > 0, \\ f_1(x) = f_2(x), & \alpha = 0, \\ f_1(x) \geq f_2(x), & \alpha < 0, \end{cases}$$

and

$$f_1(x) \overset{\text{sgn}(\alpha)}{\geq} f_2(x) \Leftrightarrow \begin{cases} f_1(x) \geq f_2(x), & \alpha > 0, \\ f_1(x) = f_2(x), & \alpha = 0, \\ f_1(x) \leq f_2(x), & \alpha < 0. \end{cases}$$

### 2.1. Definition of the U-function

It is typically challenging to locate the root  $\theta^*$  of the nonlinear equation  $h(\theta) = 0$  directly. The US algorithm aims to create an alternative function  $U(\theta|\theta^{(t)})$  to replace  $h(\theta)$ , transforming the challenge of solving complex nonlinear equations into solving the equation  $U(\theta|\theta^{(t)})$  with explicit solutions. First, we assume that

$$h(\theta) < 0, \quad \forall \theta > \theta^* \text{ and } h(\theta) > 0, \quad \forall \theta < \theta^*. \quad (2.2)$$

If  $h(\theta) < 0$  when  $\theta < \theta^*$ , then both sides of the equation  $h(\theta) = 0$  can be multiplied by -1, which can also obtain the same root  $\theta^*$  satisfying the assumption (2.2). Let  $\theta^{(t)}$  represent the solution after the  $(t-1)$ -th iteration, and the function  $U(\theta|\theta^{(t)})$  satisfying the following criteria is designated as the U-function of  $h(\theta)$  at  $\theta = \theta^{(t)}$ :

$$\begin{aligned} h(\theta) &\leq U(\theta|\theta^{(t)}), & \theta < \theta^{(t)}, \\ h(\theta^{(t)}) &= U(\theta^{(t)}|\theta^{(t)}), & \theta = \theta^{(t)}, \\ h(\theta) &\geq U(\theta|\theta^{(t)}), & \theta > \theta^{(t)}. \end{aligned} \quad (2.3)$$

According to the definition of the CD inequalities symbol, the above condition may be represented as

$$h(\theta) \overset{\text{sgn}(\theta - \theta^{(t)})}{\geq} U(\theta|\theta^{(t)}), \quad \forall \theta, \theta^{(t)} \in \Theta. \quad (2.4)$$

### 2.2. The U-equation

As described above, the US algorithm is an iterative approach for solving nonlinear equations, with each iteration including a U-step and an S-step. The purpose of the U-step is to find a U-function that satisfies the condition (2.4), whereas the S-step involves solving the simplified U-equation:  $U(\theta|\theta^{(t)}) = 0$  to obtain its root  $\theta^{(t+1)}$ ,

$$\theta^{(t+1)} = \text{sol} \{ U(\theta|\theta^{(t)}) = 0, \quad \forall \theta, \theta^{(t)} \in \Theta \}. \quad (2.5)$$

In typical scenarios,  $\theta^{(t+1)}$  can be explicitly expressed, even as a linear equation. Through the iterative execution of these two steps,  $\{\theta^{(t)}\}_{t=0}^{\infty}$  can gradually converge to the real root  $\theta^*$  of the U-equation.

### 2.3. The first-derivative lower bound function method

There are numerous  $U$ -functions for a given objective function  $h(\theta)$ ; as Eq (2.4) illustrates, distinct  $U$ -functions correlate to distinct US algorithms. We may express the  $U$ -function using the lower-order derivatives of the goal function  $h(\theta)$ . This can be accomplished by a variety of techniques, such as the first-derivative lower bound (FLB), second-derivative lower-upper bound (SLUB) constants method, and third-derivative lower bound (TLB) constant method [17]. These three methodologies enhance efficient solutions when the objective function is complex and the solution is not closed, each with a distinct convergence speed. In terms of maximizing the objective function, the US algorithm based on the FLB approach shares qualities with the EM algorithm and the MM algorithm, both of which exhibit linear convergence. The FLB function technique, which is dependent on the target function's first-order derivative, is mostly used in this article to generate the required  $U$ -function. First for parameter space  $\Theta$ , we suppose that there exists a certain first-derivative lower bound function  $b(\theta)$  for the first derivative of  $h(\theta)$ , i.e.,

$$h'(\theta) \geq b(\theta), \quad \forall \theta \in \Theta. \quad (2.6)$$

The  $U$ -function of  $h(\theta)$  at  $\theta = \theta^{(t)}$  can be formally defined as follows

$$U(\theta|\theta^{(t)}) \triangleq h(\theta^{(t)}) + \int_{\theta^{(t)}}^{\theta} b(z)dz, \quad \forall \theta, \theta^{(t)} \in \Theta. \quad (2.7)$$

In fact,

$$\begin{aligned} h(\theta) - U(\theta|\theta^{(t)}) &= [h(\theta) - h(\theta^{(t)})] - \int_{\theta^{(t)}}^{\theta} b(z)dz = \int_{\theta^{(t)}}^{\theta} h'(z)dz - \int_{\theta^{(t)}}^{\theta} b(z)dz \\ &= \int_{\theta^{(t)}}^{\theta} [h'(z) - b(z)]dz \stackrel{\text{sgn}(\theta - \theta^{(t)})}{\geq} 0, \quad \forall \theta, \theta^{(t)} \in \Theta. \end{aligned}$$

Let  $\theta^*$  be the unique root of the equation  $h(\theta) = 0$ , and then the corresponding US iteration is as follows:

$$\theta^{(t+1)} = \text{sol} \left\{ U(\theta|\theta^{(t)}) = h(\theta^{(t)}) + \int_{\theta^{(t)}}^{\theta} b(z)dz = 0, \forall \theta, \theta^{(t)} > 0 \right\} \triangleq g(\theta^{(t)}), \quad (2.8)$$

where  $g(\theta^{(t)}) = g(\theta^*) + (\theta^{(t)} - \theta^*)h'(\theta^*) + 0.5(\theta^{(t)} - \theta^*)^2h''(\hat{\theta})$  is the first-order Taylor expansion around  $\theta^*$ , and  $\hat{\theta}$  is a point between  $\theta^{(t)}$  and  $\theta^*$ .

Although Li and Tian [17] proposed the idea of the US algorithm, in practical applications, only the distribution of univariate and binary parameters were studied. In the case of binary parameter distribution, when discussing the solution of the scoring equation, only one parameter can be explicitly expressed with another parameter. However, when one parameter cannot be explicitly expressed by the other parameter for this more general and complex situation, whether the US algorithm can be effectively applied is not further discussed, which is the issue to be carried out in this article. For the parameters of interest, the new FLB functions are constructed in this article, then, starting with initial values, the iterative values are updated using the corresponding S-step until the convergence criteria are met.

### 3. The US algorithm for the MLE of the Type I discrete Weibull distribution

The discrete Weibull distribution's score function has a complex double exponential form, which makes it impossible to depict its solution and, thus, prevents its two parameters from being mutually expressed. For investigating the US algorithm's applicability in complicated models, we combine the US algorithm with the FLB method in this section to optimize the maximum likelihood estimation.

#### 3.1. The Type I discrete Weibull distribution

Assuming a random variable following the Weibull distribution  $W(\lambda, \beta)$ , where  $\lambda > 0$  and  $\beta > 0$ , the cumulative distribution function (CDF) of the Weibull distribution is defined as  $H(t, \lambda, \beta) = 1 - e^{-\lambda t^\beta}$ , where  $t > 0$ . Define  $\alpha = e^{-\lambda}$ , and then  $0 < \alpha < 1$ . If the probability mass function (PMF) for a random variable  $X$  can be represented as

$$P(X = x; \alpha, \beta) = \alpha^{([x]-1)^\beta} - \alpha^{[x]^\beta}, (x \geq 1),$$

then we say that  $X$  follows the Type I discrete Weibull distribution, denoted as  $X \sim DW(\theta)$ . Here,  $[x]$  represents the maximum integer less than or equal to  $x$ . When  $\beta = 1$ , the discrete Weibull distribution degenerates to the geometric distribution  $Geo(q)$  with  $q = 1 - \alpha$ .

Naturally, the cumulative distribution function of  $X$  takes the following form:

$$F(x; \alpha, \beta) = 1 - \alpha^{[x]^\beta}.$$

#### 3.2. The US algorithm for the MLE of $\alpha$ and $\beta$

This section will go into detail on the application of the US algorithm for maximum likelihood estimation of the Type I discrete Weibull distribution. Assume  $X$  is a random variable with the Type I discrete Weibull distribution  $DW(\theta)$ , where the parameter vector  $\theta = (\alpha, \beta)^T$  is in the parameter space  $\Theta \subset \mathbb{R}^2$ . Let  $\mathbf{x} = (x_1, \dots, x_n)$  denote the observed values of the random sample  $(X_1, \dots, X_n)$ . Then, the log-likelihood function of the parameter vector  $\theta$  is given by

$$\ell(\theta|\mathbf{x}) = \sum_{i=1}^n \log(\alpha^{(x_i-1)^\beta} - \alpha^{x_i^\beta}).$$

First, the first-order partial derivative of  $\ell(\theta|\mathbf{x})$  with respect to  $\alpha$  can be calculated as

$$\frac{\partial \ell(\theta|\mathbf{x})}{\partial \alpha} = \sum_{i=1}^n \alpha^{-1} \left[ \frac{\alpha^{(x_i-1)^\beta} (x_i-1)^\beta - \alpha^{x_i^\beta} x_i^\beta}{\alpha^{(x_i-1)^\beta} - \alpha^{x_i^\beta}} \right] \triangleq h_1(\alpha). \quad (3.1)$$

Next, we construct the FLB function regarding the parameter  $\alpha$ :

$$\begin{aligned}
 h'_1(\alpha) &= \sum_{i=1}^n \frac{[(x_i - 1)^\beta - 1](x_i - 1)^\beta \alpha^{(x_i-1)^\beta-2} - (x_i^\beta - 1)x_i^\beta \alpha^{x_i^\beta-2}}{(\alpha^{(x_i-1)^\beta} - \alpha^{x_i^\beta})^2} (\alpha^{(x_i-1)^\beta} - \alpha^{x_i^\beta}) \\
 &\quad - \sum_{i=1}^n \frac{[\alpha^{(x_i-1)^\beta} (x_i - 1)^\beta - \alpha^{x_i^\beta-1} x_i^\beta] ((x_i - 1)^\beta \alpha^{(x_i-1)^\beta-1} - x_i^\beta \alpha^{x_i^\beta-1})}{(\alpha^{(x_i-1)^\beta} - \alpha^{x_i^\beta})^2} \\
 &\geq - \sum_{i=1}^n \frac{\left[ \alpha^{2(x_i-1)^\beta-2} [(x_i - 1)^\beta + (x_i - 1)^{2\beta}] + \alpha^{x_i^\beta+(x_i-1)^\beta-2} [(x_i - 1)^{2\beta} + x_i^{2\beta}] + \alpha^{2x_i^\beta-2} [x_i^{2\beta+x_i^\beta}] \right]}{(\alpha^{(x_i-1)^\beta} - \alpha^{x_i^\beta})} \quad (3.2) \\
 &\geq - \sum_{i=1}^n \alpha^{2(x_i-1)^\beta-2} \frac{[2(x_i - 1)^{2\beta} + 2x_i^{2\beta} + (x_i - 1)^\beta + x_i^\beta]}{(\alpha^{(x_i-1)^\beta} - \alpha^{x_i^\beta})^2} \\
 &\geq - \sum_{i=1}^n \frac{4x_i^{2\beta} + 2x_i^\beta}{(\alpha - \alpha^2)^2} = - \sum_{i=1}^n 3[4x_i^{2\beta} + 2x_i^\beta] \left[ \frac{1}{\alpha^2} + \frac{1}{(1 - \alpha)^2} \right] \triangleq b_1(\alpha).
 \end{aligned}$$

Then, the US iteration of  $\alpha$  can be obtained as follows:

$$\begin{aligned}
 \alpha^{(t+1)} &= \text{sol} \left[ U_\alpha(\alpha, \beta | \alpha^{(t)}, \beta^{(t)}) = h_1(\alpha^{(t)}) + \int_{\alpha^{(t)}}^{\alpha} b_1(z) dz = 0 \right] \\
 &= \text{sol} \left[ C_1 + 3C_2 \left[ \frac{1}{\alpha} - \frac{1}{1 - \alpha} \right] - 3C_2 \left[ \frac{1}{\alpha^{(t)}} - \frac{1}{1 - \alpha^{(t)}} \right] = 0 \right] \quad (3.3) \\
 &= \text{sol} \left[ (C_3 - C_1)\alpha^2 - (6C_2 + C_3 - C_1)\alpha + 3C_2 = 0 \right],
 \end{aligned}$$

where  $C_1 = h_1(\alpha^{(t)})$ ,  $C_2 = \sum_{i=1}^n [4x_i^{2\beta^{(t)}} + 2x_i^{\beta^{(t)}}]$ , and  $C_3 = 3C_2 \left[ \frac{1}{\alpha^{(t)}} - \frac{1}{1 - \alpha^{(t)}} \right]$ . Similarly, we can obtain the first-order partial derivative of  $\ell(\theta | \mathbf{x})$  with respect to  $\beta$ ,

$$\frac{\partial \ell(\theta | \mathbf{x})}{\partial \beta} = \sum_{i=1}^n \log \alpha \left[ \frac{\alpha^{(x_i-1)^\beta} (x_i - 1)^\beta \log(x_i - 1) - \alpha^{x_i^\beta} x_i^\beta \log(x_i)}{\alpha^{(x_i-1)^\beta} - \alpha^{x_i^\beta}} \right] \triangleq h_2(\beta). \quad (3.4)$$

In order to construct the FLB function and derive the US algorithm without explicit solutions for the two parameters, we need the following two lemmas.

**Lemma 1.** [18] Given that  $\theta > 0$ , we have

$$-e^{-\theta} \geq -\frac{4e^{2\max(\theta^{(t)}-1, 0)}}{(2\theta^{(t)} - \theta)^2}, \quad \forall 0 \leq \theta \leq 2\theta^{(t)} \quad \text{and} \quad \theta^{(t)} > 0.$$

**Lemma 2.** [18] For any  $\theta \geq 0$ , we have

$$-e^{-\theta} \geq -\frac{2}{3}\theta^{-2}.$$

We will then build the FLB function with respect to  $\beta$ . First, we calculate  $h_2(\beta)$ 's first derivative.

$$\begin{aligned} h_2'(\beta) &= \sum_{i=1}^n \log(\alpha) \frac{\left[ \log^2(x_i - 1)(x_i - 1)^{2\beta} \alpha^{2(x_i-1)\beta} (1 - \log \alpha) + \log^2(x_i) x_i^{2\beta} \alpha^{2(x_i-1)\beta} (1 - \log(\alpha)) \right]}{(\alpha^{(x_i-1)\beta} - \alpha^{x_i^\beta})^2} \\ &+ \sum_{i=1}^n \log(\alpha) \frac{\left[ 2 \log(\alpha) \log(x_i) \log(x_i - 1) x_i^\beta (x_i - 1)^\beta - \log^2(x_i - 1)(x_i - 1)^\beta - \log^2(x_i) x_i^\beta \right] \alpha^{2x_i^\beta}}{(\alpha^{(x_i-1)\beta} - \alpha^{x_i^\beta})^2} \\ &\geq \sum_{i=1}^n \frac{\left[ 2 \log^2(x_i) x_i^{2\beta} \alpha^{2(x_i-1)\beta} [1 - \log(\alpha)] - 2 \log^2(x_i - 1)(x_i - 1)^\beta \alpha^{2x_i^\beta} [1 - \log(\alpha)] \right]}{(\alpha^{(x_i-1)\beta} - \alpha^{x_i^\beta})} \\ &\geq \sum_{i=1}^n \frac{2 \log(\alpha)(1 - \log(\alpha)) \log^2(x_i) x_i^{2\beta} \alpha^{2(x_i-1)\beta}}{(\alpha^{(x_i-1)\beta} - \alpha^{x_i^\beta})^2} \\ &\geq -\frac{2 \log(\alpha)(1 - \log(\alpha))}{[1 - \alpha]^2} \left[ -\sum_{i=1}^n \log^2(x_i) x_i^{2\beta} \right]. \end{aligned}$$

It can be deduced from Lemma 1 and Lemma 2 that

$$\begin{aligned} -\sum_{i=1}^n \log^2(x_i) x_i^{2\beta} &= -\sum_{i=1}^n e^{2\beta \log(x_i)} \log^2(x_i) \\ &\geq -\sum_{i=1}^n \left[ \frac{4e^{2\max(2\beta^{(t)} \log(x_i) - 1, 0)} I(\log(x_i) > 0)}{[4\beta^{(t)} \log(x_i) - 2\beta]^2} + \frac{2I(\log(x_i) \leq 0)}{3[-2\beta \log(x_i)]^2} \right] \log^2(x_i) \\ &= -\sum_{i=1}^n \left[ \frac{4e^{2\max(2\beta^{(t)} \log(x_i) - 1, 0)} I(x_i > 1)}{(4\beta^{(t)} - 2\beta)^2} + \frac{I(x_i = 1)}{6\beta^2} \right], \end{aligned}$$

where  $I(\cdot)$  is the indicator function. Then, the corresponding FLB function is obtained as follows:

$$\frac{2 \log(\alpha)(1 - \log(\alpha))}{[1 - \alpha]^2} \sum_{i=1}^n \left[ \frac{e^{2\max(2\beta^{(t)} \log(x_i) - 1, 0)} I(x_i > 1)}{(2\beta^{(t)} - \beta)^2} + \frac{I(x_i = 1)}{6\beta^2} \right] \triangleq b_2(\beta). \quad (3.5)$$

Therefore, the US iterative process for parameter  $\beta$  can be given by

$$\begin{aligned} \beta^{(t+1)} &= \text{sol} \left[ U_\beta(\alpha, \beta | \alpha^{(t+1)}, \beta^{(t)}) = h_2(\beta^{(t)}) + \int_{\beta^{(t)}}^{\beta} b_2(z) dz = 0 \right] \\ &= \text{sol} \left[ h_2(\beta^{(t)}) + a_1 \left[ \frac{a_2}{2\beta^{(t)} - \beta} - \frac{I(x_i = 1)}{6\beta} \right] - a_1 \left[ \frac{a_2}{\beta^{(t)}} - \frac{I(x_i = 1)}{6\beta^{(t)}} \right] \right] \\ &= \text{sol} \left[ 6a_3\beta^2 - a_4\beta + a_5 = 0 \right], \end{aligned} \quad (3.6)$$



where

$$a_1 = \frac{2 \log(\alpha^{(t+1)})(1 - \log(\alpha^{(t+1)}))}{(1 - \alpha^{(t+1)})^2},$$

$$a_2 = e^{2\max(2\beta^{(t)} \log(x_i) - 1, 0)} I(x_i > 1),$$

$$a_3 = h_2(\beta^{(t)}) - a_1 \left[ \frac{a_2}{\beta^{(t)}} - \frac{I(x_i = 1)}{6\beta^{(t)}} \right],$$

$$a_4 = 12a_3\beta^{(t)} + 6a_1a_2 + a_1I(x_i = 1),$$

$$a_5 = 2a_1I(x_i = 1)\beta^{(t)}.$$

The algorithm process for estimating two parameters can be described as follows. In the first stage, we determine the FLB functions for parameters  $\alpha$  and  $\beta$  using Eqs (3.2) and (3.5), respectively. Subsequently, we set two initial values  $\alpha^{(t)}$  and  $\beta^{(t)}$ , calculate Eq (3.3) to get  $\alpha^{(t+1)}$ , and then compute  $\beta^{(t+1)}$  via Eq (3.6) using  $\alpha^{(t+1)}$  and  $\beta^{(t)}$ . If both of the estimates for the parameters satisfy the convergence criteria, then their corresponding values will be returned. Otherwise, we resume to update the iteration value and repeat the preceding steps until the two estimated parameters converge.

---

### Algorithm : Calculating the MLEs of $\alpha$ and $\beta$ via the US algorithm.

---

**Input:** The initial value  $\alpha^{(0)}$  and  $\beta^{(0)}$ ; The observed data  $X_{obs} = \{x_i\}_{i=0}^n$ ;

**Output:**  $\hat{\alpha}, \hat{\beta}$ .

**1** Select FLB function for parameters  $\alpha$  and  $\beta$ , respectively;

**2** Set initial values  $\alpha^{(t)}, \beta^{(t)}, t = 0$ ;

**3 repeat**

**4** Using  $\alpha^{(t)}$  and  $\beta^{(t)}$ , calculate  $\alpha^{(t+1)}$  based on (3.3);

**5** Using  $\alpha^{(t+1)}$  and  $\beta^{(t)}$ , calculate  $\beta^{(t+1)}$  based on (3.6), update  $t = t + 1$ ;

**6 until** convergence.

---

## 4. Simulation study

In this section, we conduct simulation studies to confirm the applicability to complex nonlinear equations and compare its performance to that of the classic Newton algorithm. First, we provide the calculation steps for parameter estimation using the Newton algorithm as follows:

$$\text{Step 1 : } \alpha^{(t+1)} = \alpha^{(t)} - \frac{\partial \ell(\theta|\mathbf{x})}{\partial \alpha}(\alpha^{(t)}, \beta^{(t)}) / \frac{\partial^2 \ell(\theta|\mathbf{x})}{\partial \alpha^2}(\alpha^{(t)}, \beta^{(t)}),$$

$$\text{Step 2 : } \beta^{(t+1)} = \beta^{(t)} - \frac{\partial \ell(\theta|\mathbf{x})}{\partial \beta}(\alpha^{(t+1)}, \beta^{(t)}) / \frac{\partial^2 \ell(\theta|\mathbf{x})}{\partial \beta^2}(\alpha^{(t+1)}, \beta^{(t)}).$$

**Table 1.** The MSE and percentage from simulated data for  $\beta = 0.5$ .

Sample size 50				
Parameters	$(\alpha, \beta) = (0.2, 0.5)$		$(\alpha, \beta) = (0.4, 0.5)$	
Algorithms	Newton	US	Newton	US
Percentage	100%	100%	100%	100%
MSE ( $\hat{\alpha}$ )	0.00312	0.00234	0.00429	0.00345
MSE ( $\hat{\beta}$ )	0.02222	0.02031	0.00931	0.00631
Parameters	$(\alpha, \beta) = (0.6, 0.5)$		$(\alpha, \beta) = (0.8, 0.5)$	
Algorithms	Newton	US	Newton	US
Percentage	100%	100%	100%	100%
MSE ( $\hat{\alpha}$ )	0.00422	0.00217	0.00233	0.00117
MSE ( $\hat{\beta}$ )	0.00535	0.00115	0.00547	0.00062
Sample size 100				
Parameters	$(\alpha, \beta) = (0.2, 0.5)$		$(\alpha, \beta) = (0.4, 0.5)$	
Algorithms	Newton	US	Newton	US
Percentage	100%	100%	100%	100%
MSE ( $\hat{\alpha}$ )	0.00149	0.00122	0.00253	0.00208
MSE ( $\hat{\beta}$ )	0.01294	0.01072	0.00451	0.00307
Parameters	$(\alpha, \beta) = (0.6, 0.5)$		$(\alpha, \beta) = (0.8, 0.5)$	
Algorithms	Newton	US	Newton	US
Percentage	100%	100%	100%	100%
MSE ( $\hat{\alpha}$ )	0.00193	0.00178	0.00106	0.00081
MSE ( $\hat{\beta}$ )	0.00215	0.00079	0.00207	0.00047
Sample size 200				
Parameters	$(\alpha, \beta) = (0.2, 0.5)$		$(\alpha, \beta) = (0.4, 0.5)$	
Algorithms	Newton	US	Newton	US
Percentage	100%	100%	100%	100%
MSE ( $\hat{\alpha}$ )	0.00082	0.00069	0.00128	0.00117
MSE ( $\hat{\beta}$ )	0.00647	0.00544	0.00244	0.00183
Parameters	$(\alpha, \beta) = (0.6, 0.5)$		$(\alpha, \beta) = (0.8, 0.5)$	
Algorithms	Newton	US	Newton	US
Percentage	100%	100%	100%	100%
MSE ( $\hat{\alpha}$ )	0.00106	0.00073	0.00064	0.00048
MSE ( $\hat{\beta}$ )	0.00112	0.00057	0.00102	0.00045

**Table 2.** The MSE and percentage from simulated data for  $\beta = 1.0$ .

Sample size 50				
Parameters	$(\alpha, \beta) = (0.2, 1.0)$		$(\alpha, \beta) = (0.4, 1.0)$	
Algorithms	Newton	US	Newton	US
Percentage	87%	100%	100%	100%
MSE ( $\hat{\alpha}$ )	0.00273	0.00214	0.00542	0.00259
MSE ( $\hat{\beta}$ )	0.06586	0.05798	0.05064	0.02234
Parameters	$(\alpha, \beta) = (0.6, 1.0)$		$(\alpha, \beta) = (0.8, 1.0)$	
Algorithms	Newton	US	Newton	US
Percentage	100%	100%	100%	100%
MSE ( $\hat{\alpha}$ )	0.00492	0.00297	0.00269	0.00090
MSE ( $\hat{\beta}$ )	0.02579	0.01286	0.01075	0.00106
Sample size 100				
Parameters	$(\alpha, \beta) = (0.2, 1.0)$		$(\alpha, \beta) = (0.4, 1.0)$	
Algorithms	Newton	US	Newton	US
Percentage	97%	100%	100%	100%
MSE ( $\hat{\alpha}$ )	0.00160	0.00114	0.00260	0.00150
MSE ( $\hat{\beta}$ )	0.04316	0.02192	0.02013	0.01119
Parameters	$(\alpha, \beta) = (0.6, 1.0)$		$(\alpha, \beta) = (0.8, 1.0)$	
Algorithms	Newton	US	Newton	US
Percentage	100%	100%	100%	100%
MSE ( $\hat{\alpha}$ )	0.00196	0.00123	0.00132	0.00065
MSE ( $\hat{\beta}$ )	0.01236	0.00674	0.00922	0.00091
Sample size 200				
Parameters	$(\alpha, \beta) = (0.2, 1.0)$		$(\alpha, \beta) = (0.4, 1.0)$	
Algorithms	Newton	US	Newton	US
Percentage	100%	100%	100%	100%
MSE ( $\hat{\alpha}$ )	0.00078	0.00048	0.00114	0.00074
MSE ( $\hat{\beta}$ )	0.01835	0.00674	0.00866	0.00607
Parameters	$(\alpha, \beta) = (0.6, 1.0)$		$(\alpha, \beta) = (0.8, 1.0)$	
Algorithms	Newton	US	Newton	US
Percentage	100%	100%	100%	100%
MSE ( $\hat{\alpha}$ )	0.00109	0.00065	0.00051	0.00030
MSE ( $\hat{\beta}$ )	0.00541	0.00320	0.00319	0.00088

**Table 3.** The MSE and percentage from simulated data for  $\beta = 1.5$ .

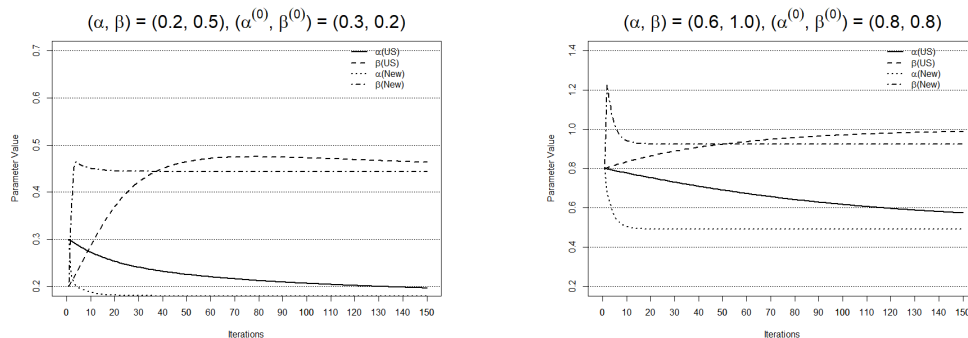
Sample size 50				
Parameters	$(\alpha, \beta) = (0.2, 1.5)$		$(\alpha, \beta) = (0.4, 1.5)$	
Algorithms	Newton	US	Newton	US
Percentage	39%	100%	98%	100%
MSE ( $\hat{\alpha}$ )	0.00529	0.00229	0.00520	0.00276
MSE ( $\hat{\beta}$ )	0.10368	0.10098	0.08433	0.05747
Parameters	$(\alpha, \beta) = (0.6, 1.5)$		$(\alpha, \beta) = (0.8, 1.5)$	
Algorithms	Newton	US	Newton	US
Percentage	100%	100%	100%	100%
MSE ( $\hat{\alpha}$ )	0.00375	0.00309	0.00334	0.00187
MSE ( $\hat{\beta}$ )	0.05022	0.03443	0.04337	0.00426
Sample size 100				
Parameters	$(\alpha, \beta) = (0.2, 1.5)$		$(\alpha, \beta) = (0.4, 1.5)$	
Algorithms	Newton	US	Newton	US
Percentage	59%	100%	100%	100%
MSE ( $\hat{\alpha}$ )	0.00141	0.00108	0.00183	0.00152
MSE ( $\hat{\beta}$ )	0.04498	0.04198	0.05027	0.02460
Parameters	$(\alpha, \beta) = (0.6, 1.5)$		$(\alpha, \beta) = (0.8, 1.5)$	
Algorithms	Newton	US	Newton	US
Percentage	100%	100%	100%	100%
MSE ( $\hat{\alpha}$ )	0.00193	0.00115	0.00123	0.00071
MSE ( $\hat{\beta}$ )	0.02331	0.00817	0.01853	0.00367
Sample size 200				
Parameters	$(\alpha, \beta) = (0.2, 1.5)$		$(\alpha, \beta) = (0.4, 1.5)$	
Algorithms	Newton	US	Newton	US
Percentage	85%	100%	100%	100%
MSE ( $\hat{\alpha}$ )	0.00074	0.00070	0.00121	0.00099
MSE ( $\hat{\beta}$ )	0.03839	0.02496	0.01860	0.00719
Parameters	$(\alpha, \beta) = (0.6, 1.5)$		$(\alpha, \beta) = (0.8, 1.5)$	
Algorithms	Newton	US	Newton	US
Percentage	100%	100%	100%	100%
MSE ( $\hat{\alpha}$ )	0.00067	0.00049	0.00065	0.00048
MSE ( $\hat{\beta}$ )	0.00921	0.00393	0.00960	0.00324

**Table 4.** The MSE and percentage from simulated data for  $\beta = 2.0$ .

Sample size 50				
Parameters	$(\alpha, \beta) = (0.2, 2.0)$		$(\alpha, \beta) = (0.4, 2.0)$	
Algorithms	Newton	US	Newton	US
Percentage	5%	100%	73%	100%
MSE ( $\hat{\alpha}$ )	0.00589	0.00235	0.00445	0.00286
MSE ( $\hat{\beta}$ )	0.50894	0.15827	0.06508	0.05573
Parameters	$(\alpha, \beta) = (0.6, 2.0)$		$(\alpha, \beta) = (0.8, 2.0)$	
Algorithms	Newton	US	Newton	US
Percentage	99%	100%	100%	100%
MSE ( $\hat{\alpha}$ )	0.00478	0.00198	0.00205	0.00153
MSE ( $\hat{\beta}$ )	0.11310	0.01770	0.06167	0.01236
Sample size 100				
Parameters	$(\alpha, \beta) = (0.2, 2.0)$		$(\alpha, \beta) = (0.4, 2.0)$	
Algorithms	Newton	US	Newton	US
Percentage	7%	100%	89%	100%
MSE ( $\hat{\alpha}$ )	0.00258	0.00155	0.00213	0.00153
MSE ( $\hat{\beta}$ )	0.02659	0.07759	0.05018	0.03908
Parameters	$(\alpha, \beta) = (0.6, 2.0)$		$(\alpha, \beta) = (0.8, 2.0)$	
Algorithms	Newton	US	Newton	US
Percentage	100%	100%	100%	100%
MSE ( $\hat{\alpha}$ )	0.00201	0.00092	0.00121	0.00084
MSE ( $\hat{\beta}$ )	0.04086	0.00654	0.03288	0.00463
Sample size 200				
Parameters	$(\alpha, \beta) = (0.2, 2.0)$		$(\alpha, \beta) = (0.4, 2.0)$	
Algorithms	Newton	US	Newton	US
Percentage	27%	100%	99%	100%
MSE ( $\hat{\alpha}$ )	0.00136	0.00077	0.00129	0.00102
MSE ( $\hat{\beta}$ )	0.10796	0.05642	0.04357	0.02541
Parameters	$(\alpha, \beta) = (0.6, 2.0)$		$(\alpha, \beta) = (0.8, 2.0)$	
Algorithms	Newton	US	Newton	US
Percentage	100%	100%	100%	100%
MSE ( $\hat{\alpha}$ )	0.00116	0.00054	0.00066	0.00041
MSE ( $\hat{\beta}$ )	0.02345	0.00357	0.01886	0.00200

The sample size of the studies is set as  $n = (50, 100, 200)$ , and the parameters are set as  $\alpha = (0.2, 0.4, 0.6, 0.8)$  and  $\beta = (0.5, 1.0, 1.5, 2.0)$ , respectively. We independently generated  $X_1^{(k)}, \dots, X_n^{(k)} \stackrel{iid}{\sim} DW(\alpha, \beta)$ , where  $k = 1, \dots, K (K = 1000)$ . The MLE of the parameters under the US algorithm were computed via Eqs (3.2) and (3.4). For every combination of parameters, we ran 1000 iterations of the experiments and evaluated the two methods' fitting performance using the convergence percentage and the mean squared error (MSE) of parameter estimation.

Tables 1–4 display the outcomes of the two algorithms' simulations for each scenario. The MSE of the parameters under both algorithms progressively drops as the sample size rises, according to the statistics in the table, suggesting that both techniques are asymptotically unbiased. When the value of  $\beta$  is fixed, as the value of  $\alpha$  increases, the MSE of  $\beta$  will steadily decrease. Overall, the MSE of both parameters under the US algorithm is smaller than that of the Newton algorithm, suggesting that the US algorithm performs better when it comes to estimation. Furthermore, it is evident from the table's convergence percentages that the US algorithm is more stable.



**Figure 1.** Simulation results of both algorithms.

The trend of the predicted values of  $\alpha$  and  $\beta$  using the US algorithm and Newton technique as the number of iterations grows is shown in Figure 1. From a stability standpoint, the Newton method shows significant instability and often requires multiple twists to preserve the correct trend, while the US algorithm approaches the true values of parameters monotonically. From a convergence speed perspective, the FLB method exhibits linear convergence, while the Newton algorithm demonstrates quadratic convergence. Consequently, the FLB method has a comparatively slower convergence rate.

## 5. Applications

This section describes the analysis of two different real data sets to illustrate the applicability of the US algorithm. The first data set in Table 5 contains the remission times in weeks of 20 leukemia patients with treatment studied by Hassan et al. [19]. Another data set is from the National Highway Traffic Safety Administration ([www-fars.nhtsa.dot.gov](http://www-fars.nhtsa.dot.gov)) of the United States, which reports the number of fatalities due to motor vehicle accidents among children under the age of 5 in 32 states during the year 2022.

**Table 5.** The leukaemia patients data and the vehicle fatalities data.

Data 1	1	3	3	6	7	7	10	12	14	15	18	19	22	26	28	29	34	40	48	49												
Data 2	15	1	6	3	23	6	1	25	14	4	13	15	7	14	2	2	8	12	3	6	12	13	2	8	2	10	5	15	47	7	3	4

We model the two data sets with the Type I discrete Weibull distribution and the geometric distribution. The fitting results for the leukemia patients data by two distributions are provided in Table 6. The results of the Cramer-von Mises test, Anderson-Darling test, and Kolmogorov-Smirnov test show that the two distributions can successfully fit the data set. The p-values from the three tests

of the Type I discrete Weibull distribution employing the US algorithm demonstrate the best fitting effect. Moreover, the values of Akaike information criterion (AIC) [20] and Bayesian information criterion (BIC) [21] also show that the DW distribution based on the US algorithm has better estimation effect.

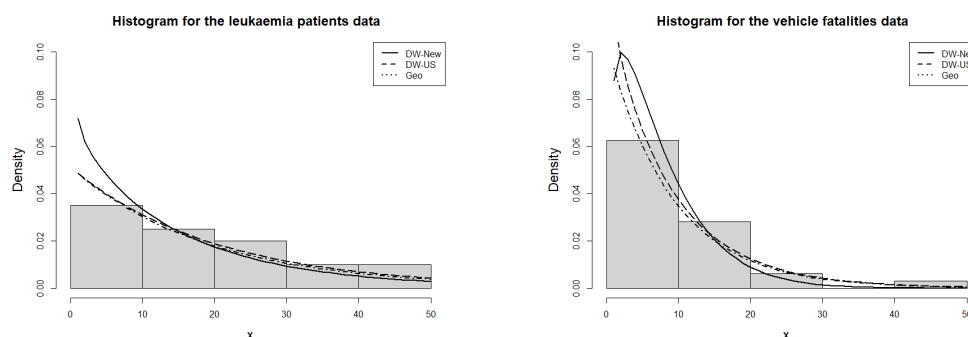
**Table 6.** Fitting for the leukaemia patients data by the DW distribution and the geometric distribution.

Methods	Estimates	AIC	BIC	p-value (KS)	p-value (AD)	p-value (CVM)
DW distribution (New)	$\alpha=0.9282 \beta=0.9477$	163.1752	165.1667	0.4207	0.1952	0.1533
DW distribution (US)	$\alpha=0.9513 \beta=1.0000$	161.9296	163.9211	0.9722	0.8817	0.8463
Geo distribution	$q=0.0512$	161.9783	162.9741	0.7035	0.4456	0.3946

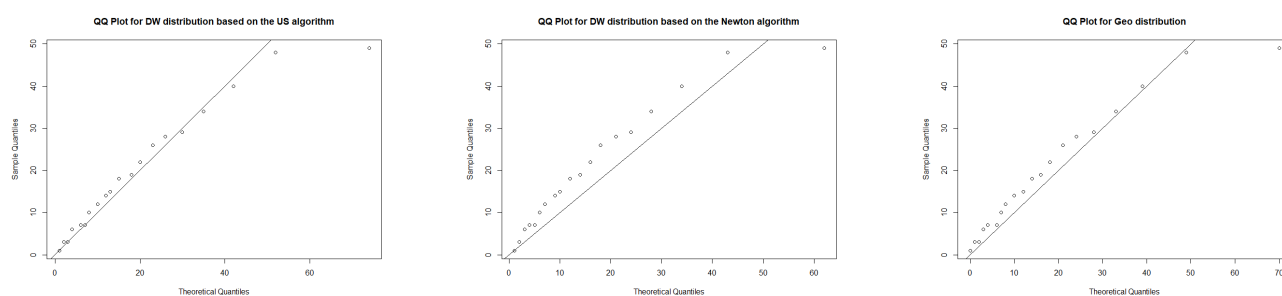
Table 7 shows the results of fitting the vehicle fatality data with the Type I discrete Weibull distribution and the geometric distribution. Comparing the p-values of the three tests at the significance level  $\alpha = 0.05$  reveals that the Geo distribution and the Type I discrete Weibull distribution estimated by the Newton algorithm are considered to be insufficient. The fitting effectiveness of the DW distribution using the US algorithm is significant, as indicated by the values of AIC and BIC. The histogram for two different data sets evaluated by the DW distribution and the Geo distribution is shown in Figure 2. Figures 3 and 4 present QQ plots for these two distributions. It is also evident that the US algorithm performs better.

**Table 7.** Fitting for the vehicle fatalities data by the DW distribution and the geometric distribution.

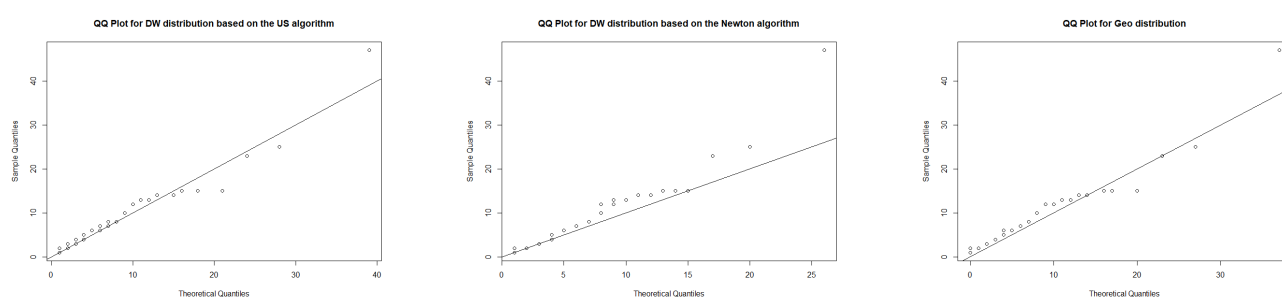
Methods	Estimates	AIC	BIC	p-value (KS)	p-value (AD)	p-value (CVM)
DW distribution (New)	$\alpha=0.9122 \beta=1.1756$	212.8066	215.2212	0.1822	0.1324	0.1755
DW distribution (US)	$\alpha=0.8852 \beta=0.9672$	209.7482	212.6796	0.4707	0.3948	0.4147
Geo distribution	$q=0.1039$	214.4938	215.9596	0.0957	0.0857	0.0941



**Figure 2.** Histogram of leukaemia patients data (left panel) and vehicle fatalities data (right panel), and the correlation density curve fitted by the DW and Geo distributions.



**Figure 3.** QQ plots of the first data for the DW (US) (left panel), DW (Newton) (middle panel), and Geo distributions (right panel).



**Figure 4.** QQ plots of the second data for the DW (US) (left panel), DW (Newton) (middle panel), and Geo distributions (right panel).

## 6. Conclusions

The US algorithm is a novel iterative method with high stability and convergence. The existing research only involves simple models such as univariate nonlinear equations or univariate functions. This paper extends the US algorithm to more complex cases of two parameter discrete distribution functions, where one parameter cannot be explicitly represented by the other parameter estimate. In order to successfully estimate the parameters, this paper combines the FLB method to perform optimization estimation of a distribution function. The simulation results for the Type I discrete Weibull distribution demonstrate that the US algorithm has good accuracy and stability. Simultaneously, for the purpose of demonstrating the applicability of the algorithm in complex situations, this paper conducted empirical research on two real data sets that follow the Type I discrete Weibull distribution, namely, the data from patients with leukemia and children who die from motor vehicle accidents. After comparing and analyzing the US method with the conventional Newton algorithm, the results show that the recommended strategy has an excellent fitting effect.

## Author contributions

Yuanhang Ouyang: Formal analysis, Writing original draft, Software, Investigation, Methodology, Data curation; Ruyun Yan: Validation, Software, Formal analysis; Jianhua Shi: Validation, Resources, Writing-review and editing, Methodology. All authors have read and approved the final version of the manuscript for publication.



## Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

## Acknowledgments

This research was conducted under a project titled “The National Social Science Fund of China” (20XTJ003).

## Conflict of interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

## References

1. M. Szymkowiak, M. Iwińska, Characterizations of discrete weibull related distributions, *Stat. Probab. Lett.*, **111** (2016), 41–48. <https://doi.org/10.1016/j.spl.2016.01.004>
2. D. Roy, Reliability measures in the discrete bivariate set-up and related characterization results for a bivariate geometric distribution, *J. Multivar. Anal.*, **46** (1993), 362–373. <https://doi.org/10.1006/jmva.1993.1065>
3. D. Roy, The discrete normal distribution, *Commun. Stat.-Theory Methods*, **32** (2003), 1871–1883.
4. H. Krishna, P. S. Pundir, Discrete burr and discrete pareto distributions, *Stat. Methodol.*, **6** (2009), 177–188. <https://doi.org/10.1016/j.stamet.2008.07.001>
5. C. Bracquemond, O. Gaudoin, A survey on discrete lifetime distributions, *Int. J. Reliab., Qual. Saf. Eng.*, **10** (2003), 69–98. <https://doi.org/10.1142/S0218539303001007>
6. K. Samuel, N. Balakrishnan, N. L. Johnson, *Continuous Multivariate Distributions: Models and Applications*, New York: John Wiley & Sons, Inc, 2000. <https://doi.org/10.1002/0471722065>
7. T. Nakagawa, S. Osaki, The discrete weibull distribution, *IEEE Trans. Reliab.*, **R-24** (1975), 300–301. <https://doi.org/10.1109/TR.1975.5214915>
8. W. E. Stein, R. Dattero, A new discrete weibull distribution, *IEEE Trans. Reliab.*, **R-33** (1984), 196–197. <https://doi.org/10.1109/TR.1984.5221777>
9. W. J. Padgett, J. D. Spurrier, On discrete failure models, *IEEE Trans. Reliab.*, **R-34** (1985), 253–256.
10. J. D. Englehardt, R. Li, The discrete weibull distribution: An alternative for correlated counts with confirmation for microbial counts in water, *Risk Anal.: Int. J.*, **31** (2011), 370–381. <https://doi.org/10.1111/j.1539-6924.2010.01520.x>
11. A. Barbiero, A comparison of methods for estimating parameters of the type i discrete weibull distribution, *Stat. Interface*, **9** (2016), 203–212. <https://doi.org/10.4310/SII.2016.v9.n2.a7>
12. A. Barbiero, Least-squares and minimum chi-square estimation in a discrete weibull model, *Commun. Stat.-Simul. Comput.*, **46** (2017), 8028–8048. <https://doi.org/10.1080/03610918.2016.1263733>

13. R. Vila, E. Y. Nakano, H. Saulo, Theoretical results on the discrete weibull distribution of nakagawa and osaki, *Statistics*, **53** (2019), 339–363. <https://doi.org/10.1080/02331888.2018.1550645>
14. H. Yoo, Application of discrete weibull regression model with multiple imputation, *Commun. Stat. Appl. Methods*, **26** (2019), 325–336. <https://doi.org/10.29220/CSAM.2019.26.3.325>
15. M. El-Morshedy, M. Eliwa, A. El-Gohary, A. Khalil, Bivariate exponentiated discrete weibull distribution: Statistical properties, estimation, simulation and applications, *Math. Sci.*, **14** (2020), 29–42.
16. P. Liu, K. C. Yuen, L. C. Wu, G. L. Tian, T. Li, Zero-one-inflated simplex regression models for the analysis of continuous proportion data, *Stat. Interface*, **13** (2020), 193–208.
17. X. Li, G.-L. Tian, The upper-crossing/solution (US) algorithm for root-finding with strongly stable convergence, 2022. <https://doi.org/10.48550/arXiv.2212.00797>
18. Y. Cai, Maximum likelihood estimates of parameters in generalized gamma distribution with self algorithm, 2023. <https://doi.org/10.48550/arXiv.2306.16419>
19. H. S. Bakouch, M. A. Jazi, S. Nadarajah, A new discrete distribution, *Statistics*, **48** (2012), 200–240. <https://doi.org/10.1080/02331888.2012.716677>
20. H. Bozdogan, Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions, *Psychometrika*, **52** (1987), 345–370. <https://doi.org/10.1007/BF02294361>
21. S. I. Vrieze, Model selection and psychological theory: A discussion of the differences between the akaike information criterion (AIC) and the bayesian information criterion (BIC), *Psychol. Methods*, **17** (2012), 228. <https://doi.org/10.1037/a0027127>



AIMS Press

©2024 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0>)