

Article

# Normal-G Class of Probability Distributions: Properties and Applications

Fábio V. J. Silveira <sup>1</sup>, Frank Gomes-Silva <sup>1,\*</sup>, Cícero C. R. Brito <sup>2</sup>, Moacyr Cunha-Filho <sup>1</sup>, Felipe R. S. Gusmão <sup>1</sup> and Sílvio F. A. Xavier-Júnior <sup>3</sup>

<sup>1</sup> Department of Statistics and Informatics, Rural Federal University of Pernambuco, Recife 52171900, Pernambuco, Brazil; fabiovjs@gmail.com (F.V.J.S.); moacyr.cunhafo@ufrpe.br (M.C.-F.); felipe556@gmail.com (F.R.S.G.)

<sup>2</sup> Federal Institute of Education, Science and Technology of Pernambuco, Recife 50740545, Pernambuco, Brazil; cicero-carlos-brito@yahoo.com.br

<sup>3</sup> Department of Statistics, Paraíba State University, Campina Grande 58429500, Paraíba, Brazil; profsilviofernando@gmail.com

\* Correspondence: franksinatrags@gmail.com

Received: 28 September 2019; Accepted: 12 November 2019; Published: 15 November 2019



**Abstract:** In this paper, we propose a novel class of probability distributions called Normal-G. It has the advantage of demanding no additional parameters besides those of the parent distribution, thereby providing parsimonious models. Furthermore, the class enjoys the property of identifiability whenever the baseline is identifiable. We present special Normal-G sub-models, which can fit asymmetrical data with either positive or negative skew. Other important mathematical properties are described, such as the series expansion of the probability density function (pdf), which is used to derive expressions for the moments and the moment generating function (mgf). We bring Monte Carlo simulation studies to investigate the behavior of the maximum likelihood estimates (MLEs) of two distributions generated by the class and we also present applications to real datasets to illustrate its usefulness.

**Keywords:** probabilistic distribution class; normal distribution; identifiability; maximum likelihood; moments

## 1. Introduction

For many purposes, statistical distributions are used in a plethora of science fields. They are regularly useful tools to describe natural and social phenomena, providing suitable models which can help dealing with real problems, such as for instance, those concerning the prediction of an event of interest. Recent works have focused attention at formulating and describing new classes of probability distributions, which are defined generally as extensions of widely known models by adding a single or more parameters to the cumulative distribution function (cdf). Hopefully, the new models will provide more flexibility and better fitting to real data. Some examples are [1,2] where a shape parameter is added to the model by exponentiating the cdf. A general method of introducing a parameter to expand a family of distributions was presented by [3]; they applied the method to create a new two-parameter extension of the exponential distribution and a new three-parameter Weibull distribution.

A natural generalization of the Normal pdf was proposed by [4] and perhaps it is the most widely known generalized Normal distribution. The power 2 appearing in the original pdf was replaced by a shape parameter  $s > 0$ . Therewith, the new pdf becomes:

$$f(x|\mu, \sigma, s) = K \exp \left\{ - \left| \frac{x - \mu}{\sigma} \right|^s \right\},$$

where  $K$  is a normalizing constant, which depends on  $\sigma$  and  $s$ . One can see that the Laplace distribution is a particular case of the generalized Normal of Nadarajah [4] when  $s = 1$ .

Azzalini [5] defined a mathematically tractable class that includes strictly (not just asymptotically) the Normal distribution. The general pdf of the class is  $2G(\lambda y)f(y)$  for  $-\infty < y < \infty$ , where  $\lambda \in \mathbb{R}$ ,  $G$  is an absolutely continuous cdf,  $\frac{d}{dy}G$  and  $f$  are pdfs symmetric about 0. Making  $G = \Phi$  and  $f = \phi$ , namely the standard normal cdf and pdf respectively, one gets to the well-known skew-normal distribution, whose pdf is  $\phi(y; \lambda) = 2\phi(y)\Phi(\lambda y)$ . It is easy to see that  $\phi(y; 0) = \phi(y)$ , but when  $\lambda \neq 0$ , the distribution is asymmetric and its coefficient of skewness has the same sign as  $\lambda$ .

A generalization denoted by compressed normal distribution was introduced by [6], whose objective was dealing with negatively skewed data (specifically with human longevity data); in this way, they induced a skew by adding  $kx$  to the denominator of the location-scale transformation, that is,

$$t(x) = \frac{x - \mu}{\sigma + kx}$$

and when  $k < 0$ , the curve presents a negative skew; for  $k > 0$ , a positive skew occurs.

Classes with one or more additional parameters usually generalize existing classes as particular cases. The McDonald-Weibull distribution [7] is an important sub-model of the McDonald class; it has three extra parameters and includes the Beta-Weibull [8] and the Kumaraswamy-Weibull [9] as special cases.

A technique to derive families of continuous distributions using a pdf as a generator was introduced by [10] and the models emerged from such method are called members of the T-X family. In other words, if  $r(t)$  is the pdf of a random variable  $T \in [a, b]$ , for  $-\infty \leq a < b \leq \infty$  and  $W(G(x))$  is a function of the cdf  $G(x)$  of a random variable  $X$  so that:

- $W(G(x)) \in [a, b]$ ;
- $W(G(x))$  is differentiable and monotonically non-decreasing;
- $W(G(x)) \rightarrow a$  as  $x \rightarrow -\infty$  and  $W(G(x)) \rightarrow b$  as  $x \rightarrow \infty$ ;

then  $F(x) = \int_a^{W(G(x))} r(t)dt$  is the cdf of a new family of distributions.

An example of a T-X family member is the Gompertz-G class [11]; to define its cdf, the chosen functions were  $W[G(x)] = -\log[1 - G(x)]$  and  $r(t) = \theta e^{\gamma t} e^{-\frac{\theta}{\gamma}(e^{\gamma t} - 1)}$  for  $t > 0$ , given that  $\theta > 0$ ,  $\gamma > 0$ . Varying  $G(x)$ , one can get different sub-models of the class.

The procedure to define a T-X family member is indeed capable to generalize a large number of distributions. Even though it can be regarded as a particular case described by the method of generating classes of probability distributions presented in the recent work of [12]. This new method has a high power of generalization. It consists of creating distribution classes by integrating a cdf, such that the limits of the integration are special functions that satisfy some conditions. Thus, the cdf of the general class is given by:

$$F(x) = \zeta(x) \sum_{j=1}^n \int_{L_j(x)}^{U_j(x)} dH(t) - \nu(x) \sum_{j=1}^n \int_{M_j(x)}^{V_j(x)} dH(t) \quad (1)$$

where  $H$  is a cdf,  $n \in \mathbb{N}$ ,  $\zeta, \nu : \mathbb{R} \mapsto \mathbb{R}$  and  $L_j, U_j, M_j, V_j : \mathbb{R} \mapsto \mathbb{R} \cup \{\pm\infty\}$  are the aforementioned special functions that will be discussed in the next section.

Based on this innovative method, we introduce the Normal-G class of distributions. We consider that this extension will yield good submodels. This paper aims to investigate and compare some of them with other competitive extended probability distributions.

## 2. The Normal-G Class and Some Mathematical Properties

The method established by [12] states that if  $H, \zeta, \nu : \mathbb{R} \mapsto \mathbb{R}$  and  $L_j, U_j, M_j, V_j : \mathbb{R} \mapsto \mathbb{R} \cup \{\pm\infty\}$  for  $j = 1, 2, 3, \dots, n$  are monotonic and right continuous functions such that:

- (c1)  $H$  is a cdf and  $\zeta$  and  $\nu$  are non-negative;

- (c2)  $\zeta(x)$ ,  $U_j(x)$  and  $M_j(x)$  are non-decreasing and  $\nu(x)$ ,  $V_j(x)$ ,  $L_j(x)$  are non-increasing  $\forall j = 1, 2, 3, \dots, n$ ;
- (c3) If  $\lim_{x \rightarrow -\infty} \zeta(x) \neq \lim_{x \rightarrow -\infty} \nu(x)$ , then  $\lim_{x \rightarrow -\infty} \zeta(x) = 0$  **or**  $\lim_{x \rightarrow -\infty} U_j(x) = \lim_{x \rightarrow -\infty} L_j(x) \forall j = 1, 2, 3, \dots, n$ , and  $\lim_{x \rightarrow -\infty} \nu(x) = 0$  **or**  $\lim_{x \rightarrow -\infty} M_j(x) = \lim_{x \rightarrow -\infty} V_j(x) \forall j = 1, 2, 3, \dots, n$ ;
- (c4) If  $\lim_{x \rightarrow -\infty} \zeta(x) = \lim_{x \rightarrow -\infty} \nu(x) \neq 0$ , then  $\lim_{x \rightarrow -\infty} U_j(x) = \lim_{x \rightarrow -\infty} V_j(x)$  and  $\lim_{x \rightarrow -\infty} M_j(x) = \lim_{x \rightarrow -\infty} L_j(x) \forall j = 1, 2, 3, \dots, n$ ;
- (c5)  $\lim_{x \rightarrow -\infty} L_j(x) \leq \lim_{x \rightarrow -\infty} U_j(x)$  and if  $\lim_{x \rightarrow -\infty} \nu(x) \neq 0$ , then  $\lim_{x \rightarrow +\infty} M_j(x) \leq \lim_{x \rightarrow +\infty} V_j(x) \forall j = 1, 2, 3, \dots, n$ ;
- (c6)  $\lim_{x \rightarrow +\infty} U_n(x) \geq \sup\{x \in \mathbb{R} : H(x) < 1\}$  and  $\lim_{x \rightarrow +\infty} L_1(x) \leq \inf\{x \in \mathbb{R} : H(x) > 0\}$ ;
- (c7)  $\lim_{x \rightarrow +\infty} \zeta(x) = 1$ ;
- (c8)  $\lim_{x \rightarrow +\infty} \nu(x) = 0$  or  $\lim_{x \rightarrow +\infty} M_j(x) = \lim_{x \rightarrow +\infty} V_j(x) \forall j = 1, 2, 3, \dots, n$  and  $n \geq 1$ ;
- (c9)  $\lim_{x \rightarrow +\infty} U_j(x) = \lim_{x \rightarrow +\infty} L_{j+1}(x) \forall j = 1, 2, 3, \dots, n - 1$  and  $n \geq 2$ ;
- (c10)  $H$  is a cdf without points of discontinuity or all functions  $L_j(x)$  and  $V_j(x)$  are constant at the right of the vicinity of points whose image are points of discontinuity of  $H$ , being also continuous in that points. Moreover,  $H$  does not have any point of discontinuity in the set  $\left\{ \lim_{x \rightarrow \pm\infty} L_j(x), \lim_{x \rightarrow \pm\infty} U_j(x), \lim_{x \rightarrow \pm\infty} M_j(x), \lim_{x \rightarrow \pm\infty} V_j(x) \right\}$  for some  $j = 1, 2, 3, \dots, n$ ;

then Equation (1) is a cdf.

Let  $n = 1$ ,  $\zeta(x) = 1$ ,  $\nu(x) = 0$ ,  $L_1(x) = -\infty$ ,  $U_1(x) = [2G(x) - 1] / (G(x)[1 - G(x)])$  and  $H(t) = \Phi(t)$ ; the function in Equation (1) turns into:

$$F_G(x) = \int_{-\infty}^{\frac{2G(x)-1}{G(x)(1-G(x))}} d\Phi(t), \quad (2)$$

where  $G(x)$  is a cdf. Since  $\nu(x) = 0$ , there is no need to specify  $M_1(x)$  and  $V_1(x)$ . The conditions (c1), (c7), (c8) and (c10) are straightforward; clearly (c4), (c5) and (c9) do not need to be verified in this case. Given that  $G(x)$  is non-decreasing:

$$\begin{aligned} x_1 < x_2 &\Rightarrow G(x_1) \leq G(x_2) \Rightarrow \frac{1}{1 - G(x_1)} \leq \frac{1}{1 - G(x_2)} \\ &\Rightarrow \frac{1}{1 - G(x_1)} - \frac{1}{G(x_1)} \leq \frac{1}{1 - G(x_2)} - \frac{1}{G(x_2)} \\ &\Rightarrow \frac{2G(x_1) - 1}{G(x_1)(1 - G(x_1))} \leq \frac{2G(x_2) - 1}{G(x_2)(1 - G(x_2))} \Rightarrow U_1(x_1) \leq U_1(x_2), \end{aligned}$$

so  $U_1(x)$  is non-decreasing, as well as  $\zeta(x)$ ; and since  $L_1(x)$  is non-increasing, (c2) is true. Considering that  $U_1(x) = 1/[1 - G(x)] - 1/G(x)$ , it is easy to verify that  $\lim_{x \rightarrow -\infty} U_1(x) = -\infty = \lim_{x \rightarrow -\infty} L_1(x)$ ; and since  $\lim_{x \rightarrow -\infty} \nu(x) = 0$ , (c3) is satisfied. The condition (c6) is also true because  $\lim_{x \rightarrow +\infty} U_1(x) = +\infty = \sup\{x \in \mathbb{R} : H(x) < 1\}$  and  $\lim_{x \rightarrow +\infty} L_1(x) = -\infty = \inf\{x \in \mathbb{R} : H(x) > 0\}$ .

Therefore, according to the method exposed above, Equation (2) is a cdf and, from now on, we will denote it by Normal-G class of probability distributions. The new cdf can be viewed as a composed function of  $G(x)$ , which will be referred as parent distribution or baseline; in agreement with [12], if the baseline is continuous (discrete), then the Normal-G will generate a continuous (discrete) distribution, whose support will be the same as  $G(x)$ . It is worth remarking that the proposed class demands no additional parameters other than the ones of the parent distribution.

Although the Normal-G class has been defined as a composed function of a single  $G(x)$ , it is possible to formulate classes that depend on more than one baseline; see [12] for further details.

We can rewrite Equation (2) as:

$$F_G(x) = \int_{-\infty}^{\frac{2G(x)-1}{G(x)[1-G(x)]}} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt, \quad (3)$$

and since  $\phi(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}$ , and  $\Phi(x) = \int_{-\infty}^x \phi(t) dt$ , we get to:

$$F_G(x) = \Phi \left( \frac{2G(x) - 1}{G(x)[1 - G(x)]} \right). \quad (4)$$

In case of continuous  $G(x)$ , we can take the derivative of Equation (4) with respect to  $x$ :

$$f_G(x) = \phi \left( \frac{2G(x) - 1}{G(x)[1 - G(x)]} \right) \frac{1 - 2G(x)[1 - G(x)]}{G(x)^2[1 - G(x)]^2} g(x). \quad (5)$$

The expression in Equation (5) is the pdf of the class Normal-G, whose hazard rate function (hrf) is given by:

$$\tau_G(x) = \frac{\phi \left( \frac{2G(x) - 1}{G(x)[1 - G(x)]} \right)}{1 - \Phi \left( \frac{2G(x) - 1}{G(x)[1 - G(x)]} \right)} \left[ \frac{1 - 2G(x)[1 - G(x)]}{G(x)^2[1 - G(x)]^2} g(x) \right].$$

Many distributions presented in the statistical literature undergo the problem of non-identifiability. One cannot assume that the parameters of a non-identifiable model will be uniquely determined from a set of observed random variables; in other words, inferences on the parameters may not be reliable. As the Theorem 1 states, the Normal-G class is exempt from this problem, whenever the parent distribution  $G$  satisfies the property of identifiability.

**Theorem 1.** *If the cdf  $F_G$  belongs to the Normal-G class and the cdf  $G$  is identifiable, then  $F_G$  is identifiable.*

**Proof of Theorem 1.** Given that  $0 < G(x|\xi_j) < 1$  for  $j = 1, 2$ , where  $\xi_j$  is a parametric vector and assuming that  $F_G(x|\xi_1) = F_G(x|\xi_2)$ , we have:

$$\Phi \left( \frac{2G(x|\xi_1) - 1}{G(x|\xi_1)[1 - G(x|\xi_1)]} \right) = \Phi \left( \frac{2G(x|\xi_2) - 1}{G(x|\xi_2)[1 - G(x|\xi_2)]} \right).$$

Since the function  $\Phi$  is injective, we can write:

$$\frac{1}{1 - G(x|\xi_1)} - \frac{1}{G(x|\xi_1)} = \frac{1}{1 - G(x|\xi_2)} - \frac{1}{G(x|\xi_2)}$$

$$\frac{G(x|\xi_1) - G(x|\xi_2)}{[1 - G(x|\xi_1)][1 - G(x|\xi_2)]} = \frac{G(x|\xi_1) - G(x|\xi_2)}{-G(x|\xi_1)G(x|\xi_2)}$$

If  $G(x|\xi_1) \neq G(x|\xi_2)$ , then:

$$[1 - G(x|\xi_1)][1 - G(x|\xi_2)] = -G(x|\xi_1)G(x|\xi_2) \quad (6)$$

The left-hand side of Equation (6) is necessarily positive for almost all  $x \in \mathbb{R}$ , whereas the right-hand side is negative, a contradiction. Thereby,  $G(x|\xi_1) = G(x|\xi_2) \Rightarrow \xi_1 = \xi_2$ .  $\square$

### 2.1. Special Normal-G Sub-Models

Here we present two distributions from the Normal-G class.

### 2.1.1. The Normal-Weibull Distribution

Weibull is one of the most used models to describe natural phenomena and failure of several kinds of components. It is extensively used in survival analysis and reliability. In recent times, many authors have focused on new extensions for it, such as [13,14]. The two-parameter Weibull cdf is given by  $G_W(x|k, \lambda) = 1 - e^{-(x/\lambda)^k}$  for  $x \geq 0$ , where  $k, \lambda > 0$ . Replacing the baseline  $G$  in Equation (4) by  $G_W$ , we get to the Normal-Weibull cdf, namely:

$$F_{NW}(x) = \Phi \left[ \frac{e^{(x/\lambda)^k} - 2}{1 - e^{-(x/\lambda)^k}} \right], \tag{7}$$

for  $x \geq 0$ . Using Equation (5) to write the corresponding pdf, we have:

$$f_{NW}(x) = \phi \left[ \frac{e^{(x/\lambda)^k} - 2}{1 - e^{-(x/\lambda)^k}} \right] \left( \frac{kx^{k-1}}{\lambda^k} \right) \frac{1 - 2 \left[ 1 - e^{-(x/\lambda)^k} \right] e^{-(x/\lambda)^k}}{e^{-(x/\lambda)^k} \left[ 1 - e^{-(x/\lambda)^k} \right]^2}. \tag{8}$$

Plots of pdf and hrf of the Normal-Weibull distribution for different values of the parameters are portrayed in Figure 1. The different shapes of the hrf curve evince the flexibility of the model. Particularly for  $k = 1$ , the Weibull distribution is equivalent to an Exponential distribution, so the hrf is constant; in contrast, the Normal-Exponential model has an increasing hrf in some left-bounded interval.

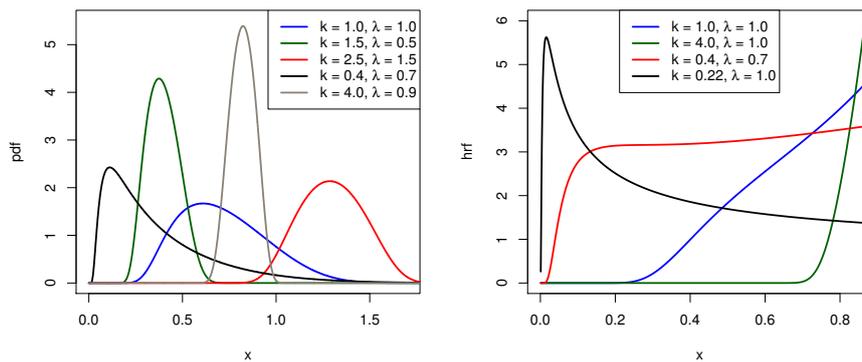


Figure 1. Plots of pdf and hrf for the Normal-Weibull distribution.

In Figure 2, the vertical axis shows the range of values of Pearson’s moment coefficient of skewness, which depends on the parameters  $k$  and  $\lambda$ . We can see in the graph that the Normal-Weibull distribution is also able to fit data with either positive or negative skew.

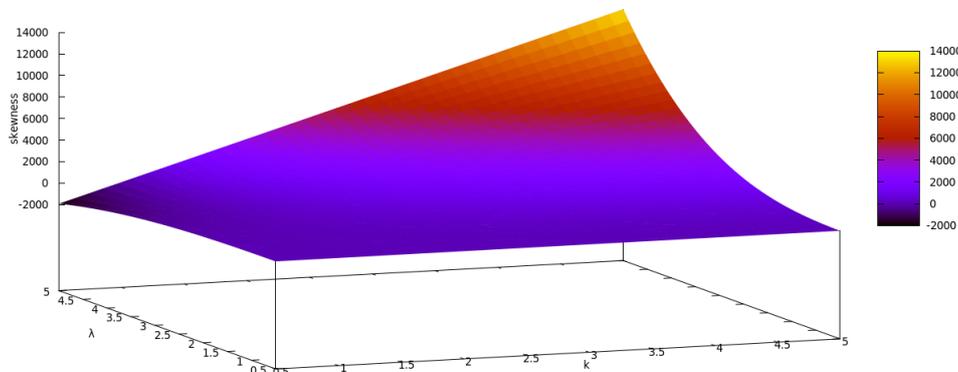


Figure 2. Skewness of the Normal-Weibull distribution.

### 2.1.2. The Normal-Log-Logistic Distribution

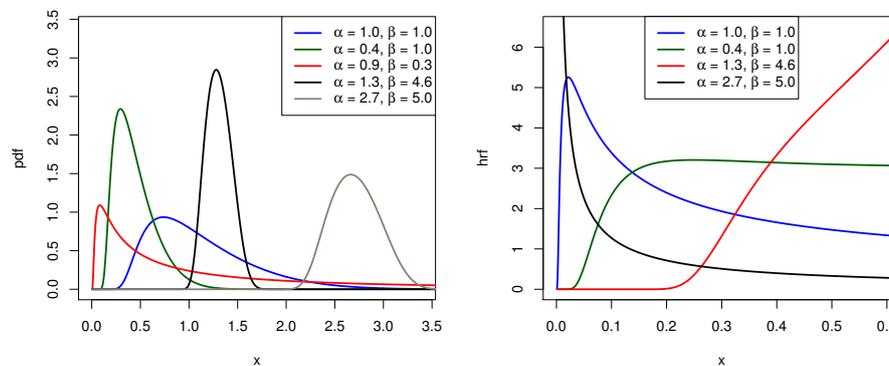
The Log-logistic distribution is commonly applied to reliability and oftentimes it works well as a lifetime model. Its cdf is given by  $G_{LL}(x|\alpha, \beta) = 1 - \left[1 + (x/\alpha)^\beta\right]^{-1}$  for  $x \geq 0$ , where  $\alpha, \beta > 0$ . The Normal-log-logistic cdf is easily obtained replacing the parent distribution  $G$  in Equation (4) by  $G_{LL}$ . Thus:

$$F_{NLL}(x) = \Phi \left[ \left(\frac{x}{\alpha}\right)^\beta - \left(\frac{x}{\alpha}\right)^{-\beta} \right], \quad (9)$$

for  $x \geq 0$ . Taking the derivative of Equation (9) with respect to  $x$ , we get to the pdf:

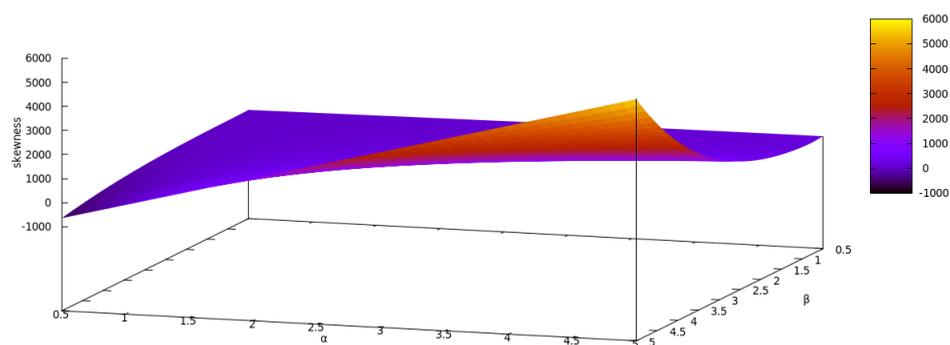
$$f_{NLL}(x) = \phi \left[ \left(\frac{x}{\alpha}\right)^\beta - \left(\frac{x}{\alpha}\right)^{-\beta} \right] \left[ 1 + \left(\frac{x}{\alpha}\right)^{2\beta} \right] \beta \alpha^\beta x^{-\beta-1}. \quad (10)$$

Figure 3 shows plots of pdf and hrf for different values of  $\alpha$  and  $\beta$ . It is worth noting that the Normal-log-logistic distribution may have a decreasing hrf of early failure. It is also possible for the hrf to be increasing or unimodal.



**Figure 3.** Plots of pdf and hrf for the Normal-log-logistic distribution.

Pearson's moment coefficient of skewness for the Normal-log-logistic distribution is depicted in Figure 4.



**Figure 4.** Skewness of the Normal-log-logistic distribution.

### 2.2. Series Representation

The normal cdf is related to the error function  $\text{erf}$  as follows:

$$\Phi(z) = \frac{1}{2} \left[ 1 + \text{erf} \left( \frac{z}{\sqrt{2}} \right) \right], \quad (11)$$

where  $\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt$ . Provided that  $\operatorname{erf}(z/\sqrt{2})$  can be linearly represented by:

$$\begin{aligned} \operatorname{erf}\left(\frac{z}{\sqrt{2}}\right) &= \frac{2}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{(-1)^n \cdot (z/\sqrt{2})^{2n+1}}{n!(2n+1)} \\ &= \sqrt{\frac{2}{\pi}} \cdot \sum_{n=0}^{\infty} \left(-\frac{1}{2}\right)^n \frac{z^{2n+1}}{n!(2n+1)}, \end{aligned} \quad (12)$$

replacing Equation (12) in Equation (11), we obtain:

$$\Phi(z) = \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \sum_{n=0}^{\infty} \left(-\frac{1}{2}\right)^n \frac{z^{2n+1}}{n!(2n+1)}. \quad (13)$$

Now, considering  $|G(x)| < 1$ , we can write:

$$\frac{2G(x) - 1}{G(x)[1 - G(x)]} = \frac{2G(x) - 1}{G(x)} \cdot \frac{1}{1 - G(x)} = \left(2 - \frac{1}{G(x)}\right) \sum_{k=0}^{\infty} G(x)^k \quad (14)$$

and replacing  $z$  of the right member of Equation (13) by the expression in Equation (14), we have:

$$\begin{aligned} \Phi\left(\frac{2G(x) - 1}{G(x)[1 - G(x)]}\right) &= \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \sum_{n=0}^{\infty} \frac{(-1/2)^n}{n!(2n+1)} \left[\left(2 - \frac{1}{G(x)}\right) \sum_{k=0}^{\infty} G(x)^k\right]^{2n+1} \\ &= \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \sum_{n=0}^{\infty} \frac{(-1/2)^n}{n!(2n+1)} \underbrace{\left(2 - \frac{1}{G(x)}\right)^{2n+1}}_{A1} \underbrace{\left[\sum_{k=0}^{\infty} G(x)^k\right]^{2n+1}}_{A2}. \end{aligned} \quad (15)$$

The right member of Equation (15) has two factors, namely, A1 and A2, that can be rewritten as power series. Concerning to A1, the binomial theorem allows us to write:

$$\begin{aligned} \left(2 - \frac{1}{G(x)}\right)^{2n+1} &= \sum_{j=0}^{2n+1} \binom{2n+1}{j} 2^{2n+1-j} \left(-\frac{1}{G(x)}\right)^j \\ &= \sum_{j=0}^{2n+1} \binom{2n+1}{j} (-1)^j \cdot 2^{2n+1-j} \cdot G(x)^{-j} \\ &= \sum_{j=0}^{2n+1} \delta_j \cdot G(x)^{-j} \end{aligned} \quad (16)$$

It is a known result related to power series raised to powers that:

$$\left[\sum_{k=0}^{\infty} a_k G(x)^k\right]^N = \sum_{k=0}^{\infty} c_k G(x)^k, \quad (17)$$

where  $c_0 = a_0^N$ ,  $c_k = \frac{1}{ka_0} \sum_{s=1}^k (sN - k + s) a_s c_{k-s}$  for  $k \geq 1$  and  $N \in \mathbb{N}$ . Setting  $N = 2n + 1$  and  $a_k = 1$  for all  $k \geq 0$ , we get to the expression A2 in Equation (15) and we can use the result in Equation (17) to write as follows:

$$\left[\sum_{k=0}^{\infty} G(x)^k\right]^{2n+1} = \sum_{k=0}^{\infty} c_k \cdot G(x)^k, \quad (18)$$

such that  $c_0 = 1$ ,  $c_k = \frac{1}{k} \sum_{s=1}^k (s[2n + 1] - k + s)c_{k-s}$  for  $k \geq 1$  and  $2n + 1 \in \mathbb{N}$ . Now replacing A1 and A2 of the Equation (15) by the right members of the Equations (16) and (18) respectively, we obtain the result below:

$$\begin{aligned} \Phi\left(\frac{2G(x) - 1}{G(x)[1 - G(x)]}\right) &= \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \sum_{n=0}^{\infty} \frac{(-1/2)^n}{n!(2n + 1)} \cdot \sum_{j=0}^{2n+1} \delta_j \cdot G(x)^{-j} \cdot \sum_{k=0}^{\infty} c_k \cdot G(x)^k \\ &= \frac{1}{2} + \sum_{n=0}^{\infty} \sum_{j=0}^{2n+1} \sum_{k=0}^{\infty} \underbrace{\binom{2n + 1}{j} \frac{(-1)^{n+j} \cdot 2^{n+1-j}}{n!(2n + 1)\sqrt{2\pi}} c_k}_{\eta_{j,n,k}} \cdot G(x)^{k-j} \\ &= \frac{1}{2} + \sum_{n,k=0}^{\infty} \sum_{j=0}^{2n+1} \eta_{j,n,k} \cdot G(x)^{k-j}. \end{aligned} \tag{19}$$

The Fubini’s theorem on differentiation allows us to write the derivative of Equation (19) with respect to  $x$  as follows:

$$f_G(x) = \sum_{n,k=0}^{\infty} \sum_{j=0}^{2n+1} \eta_{j,n,k} \cdot \underbrace{(k - j)g(x)G(x)^{k-j-1}}_{g_{k-j}(x)}. \tag{20}$$

Since  $g_{k-j}(x)$  is the pdf of a random variable of the exponentiated family, as described in [15,16], one can say that (20) is the Normal-G pdf (5) expressed as a linear combination of pdfs of exponentiated distributions. Such useful property is typically found and detailed in works on new classes of distributions; see for instance: [17–20].

### 2.3. Quantile Function

By inverting Equation (4), the quantile function associated with the Normal-G class is obtained. For simplification, let us write  $v = F_G(x)$ . From Equation (4) we have:

$$\Phi^{-1}(v) = \frac{2G(x) - 1}{G(x)[1 - G(x)]} \Rightarrow \Phi^{-1}(v)G(x)^2 + [2 - \Phi^{-1}(v)]G(x) - 1 = 0,$$

that is, a quadratic equation for  $G(x)$ , that admits the following two solutions:

$$\frac{\Phi^{-1}(v) - 2 - \sqrt{4 + \Phi^{-1}(v)^2}}{2\Phi^{-1}(v)} \quad \text{and} \quad \frac{\Phi^{-1}(v) - 2 + \sqrt{4 + \Phi^{-1}(v)^2}}{2\Phi^{-1}(v)}.$$

If the first solution above is picked, then  $G(x)$  might assume values lesser than 0 (see  $v = 0.95$  for example). On the other hand, the second one allows us to verify that  $0 < G(x) < 1$  is valid for all  $x \in \mathbb{R}$ . Finally, we can write the quantile function of Equation (4) as follows:

$$Q_F(v) = Q_G \left[ \frac{\Phi^{-1}(v) - 2 + \sqrt{4 + \Phi^{-1}(v)^2}}{2\Phi^{-1}(v)} \right], \tag{21}$$

such that  $Q_G(\cdot)$  is the quantile function of the baseline  $G$ . A uniform random number generator and (21) make the simulation of random variables following (3) quite simple. Namely, if  $Z \sim \mathcal{U}(0, 1)$ , then  $Q_F(Z)$  follows a Normal-G distribution.

#### 2.4. Raw Moments, Incomplete Moments and Moment Generating Function

Provided that  $X$  follows a Normal-G distribution, the  $r$ th raw moment of  $X$  is  $E(X^r) = \int_{-\infty}^{\infty} x^r f_G(x) dx$ , where  $f_G(x)$  is given in Equation (20) and  $r \in \mathbb{Z}_+^*$ . Using Fubini's theorem to change the order of integration and series, we have:

$$E(X^r) = \sum_{n,k=0}^{\infty} \sum_{j=0}^{2n+1} \eta_{j,n,k} \int_{-\infty}^{\infty} x^r g_{k-j}(x) dx \quad (22)$$

$$= \sum_{n,k=0}^{\infty} \sum_{j=0}^{2n+1} \eta_{j,n,k} E(Y_{k-j}^r) \quad (23)$$

where  $Y_{k-j}$  follows the exponentiated distribution whose pdf is  $g_{k-j}(x)$  shown in Equation (20).

Despite the upper infinity limit in the sums, expressions like Equation (23) are not intractable. According to [21], one can get fairly accurate results truncating each infinite sum by 20; they used numerical routines to compute accurately similar expressions for the moments of some Kumaraswamy generalized distributions.

The  $r$ th moment can also be represented in terms of the quantile function of the baseline. Defining  $u = G_{k-j}(x)$  and replacing  $x$  in Equation (22) by  $Q_G(u^{1/(k-j)})$ , we have:

$$E(X^r) = \sum_{n,k=0}^{\infty} \sum_{j=0}^{2n+1} \eta_{j,n,k} \int_0^1 \left[ Q_G \left( u^{\frac{1}{k-j}} \right) \right]^r du.$$

The  $r$ th incomplete moment of  $X$  is given by the following expression:

$$T_r(z) = \int_{-\infty}^z x^r f_G(x) dx = \sum_{n,k=0}^{\infty} \sum_{j=0}^{2n+1} \eta_{j,n,k} T_r^*(z) \quad (24)$$

where  $T_r^*(z)$  is the  $r$ th incomplete moment of  $Y_{k-j}$ . One can also write Equation (24) in terms of the quantile function of  $G$ :

$$T_r(z) = \sum_{n,k=0}^{\infty} \sum_{j=0}^{2n+1} \eta_{j,n,k} \int_0^{[G(x)]^{k-j}} \left[ Q_G \left( u^{\frac{1}{k-j}} \right) \right]^r du.$$

The mgf is a function associated with a random variable, whose moments can be straightforwardly derived using it. It is also useful to check whether two functions of random variables are equal since there is a bijection between pdfs and mgfs (when they exist). The mgf  $M_X(t)$  of  $X$  is the expected value of  $e^{tX}$ , where  $t \in (-\iota, \iota)$ ,  $\iota > 0$ . Given that  $M_{Y_{k-j}}(t)$  is the mgf of  $Y_{k-j}$ , on the lines of Equation (23), we can write:

$$\begin{aligned} M_X(t) &= \sum_{n,k=0}^{\infty} \sum_{j=0}^{2n+1} \eta_{j,n,k} \int_{-\infty}^{\infty} e^{tx} g_{k-j}(x) dx \\ &= \sum_{n,k=0}^{\infty} \sum_{j=0}^{2n+1} \eta_{j,n,k} M_{Y_{k-j}}(t). \end{aligned}$$

#### 2.5. Estimation and Inference

Attractive asymptotic properties, such as efficiency and consistency, are some of the reasons that make the maximum likelihood method the most usually applied method of parametric point estimation. The MLEs are the points that maximize the likelihood function over the domain of the parameter space. Since the logarithmic function is increasing, performing the maximization of the log-likelihood function, besides being a more convenient task, also provides the MLEs.

Given that  $\xi = (\xi_1, \dots, \xi_r)^\top$  is the  $r \times 1$  parametric vector of a random variable  $X$  that follows a Normal-G distribution,  $G(x|\xi) = G_\xi(x)$  is the baseline,  $g(x|\xi) = g_\xi(x)$  is its corresponding pdf and  $\mathbf{X} = (x_1, \dots, x_m)$  is a complete random sample of size  $m$  from  $X$ , then the log-likelihood function is:

$$\begin{aligned} \ell(\xi|\mathbf{X}) = & \sum_{j=1}^m \log \phi \left( \frac{2G_\xi(x_j) - 1}{G_\xi(x_j) [1 - G_\xi(x_j)]} \right) + \sum_{j=1}^m \log \left( 1 - 2G_\xi(x_j) + 2G_\xi^2(x_j) \right) \\ & - 2 \sum_{j=1}^m \log G_\xi(x_j) - 2 \sum_{j=1}^m \log [1 - G_\xi(x_j)] + \sum_{j=1}^m \log g_\xi(x_j). \end{aligned} \quad (25)$$

Thanks to powerful functions available within the software for statistical computing, it is possible to use numerical methods to maximize (25); for this purpose, R [22] brings the function `optim` in package `stats`.

The MLEs can also be obtained by solving the system of equations  $U(\xi|\mathbf{X}) = \mathbf{0}_r$ , where  $U(\xi|\mathbf{X}) = \nabla_\xi \ell(\xi|\mathbf{X}) = (u_i)_{1 \leq i \leq r}$  is the score vector, such that:

$$\begin{aligned} u_i = & \sum_{j=1}^m \frac{[G_\xi(x_j) - 1]^4 - G_\xi^4(x_j)}{G_\xi^3(x_j) [1 - G_\xi(x_j)]^3} \cdot \frac{\partial}{\partial \xi_i} G_\xi(x_j) + \sum_{j=1}^m \frac{4G_\xi(x_j) - 2}{1 - 2G_\xi(x_j) + 2G_\xi^2(x_j)} \cdot \frac{\partial}{\partial \xi_i} G_\xi(x_j) \\ & - 2 \sum_{j=1}^m \frac{1}{G_\xi(x_j)} \cdot \frac{\partial}{\partial \xi_i} G_\xi(x_j) + 2 \sum_{j=1}^m \frac{1}{1 - G_\xi(x_j)} \cdot \frac{\partial}{\partial \xi_i} G_\xi(x_j) + \sum_{j=1}^m \frac{1}{g_\xi(x_j)} \cdot \frac{\partial}{\partial \xi_i} g_\xi(x_j) \end{aligned}$$

and  $\mathbf{0}_r$  is a  $r \times 1$  vector of zeros.

The information matrix  $J(\xi|\mathbf{X})$  is essential to construct confidence intervals and to test hypotheses on  $\xi$ . The expectation of  $J(\xi|\mathbf{X})$  is the expected Fisher information matrix  $\mathcal{I}_\xi$  and under certain conditions of regularity,  $\sqrt{m}(\hat{\xi} - \xi)$  follows approximately a multivariate normal distribution  $N_r(\mathbf{0}_r, \mathcal{I}_\xi^{-1})$ . The expression for  $J(\xi|\mathbf{X})$  is presented in Appendix A.

### 3. Numerical Analysis

#### 3.1. Simulation Study

We used the free software R version 3.4.4 [22] to carry out the Monte Carlo simulation study; the number of replications was 10,000. The pseudo-random samples were generated via Von Neumann's acceptance-rejection method [23]. This simple procedure requires the corresponding pdf  $y = f(x)$ , a minorant and a majorant for  $x$  and a majorant for  $y$ ; it is not necessary to implement the quantile function in this case. Four sample sizes, namely  $n = 50, 100, 200$  and  $500$ , and five different values for the vector of parameters were considered. For each scenario, we calculated the bias and the mean squared error (MSE) as follows:

$$\text{Bias}_i = \frac{1}{10000} \sum_{j=1}^{10000} (\hat{\xi}_{ij} - \xi_i), \quad \text{MSE}_i = \frac{1}{10000} \sum_{j=1}^{10000} (\hat{\xi}_{ij} - \xi_i)^2$$

where  $\xi_i$  is the  $i$ -th element of the vector of parameters  $\xi = (\xi_1, \dots, \xi_r)^\top$  and  $\hat{\xi}_{ij}$  is the estimate for  $\xi_i$  at the  $j$ -th replication. The log-likelihood function was maximized using the technique of simulated annealing, available by the `optim` subroutine, for which the user has to pass a vector  $\xi_0$  of initial values. At first, we took  $\xi_0 = \mathbf{1}_r$ , namely a  $r \times 1$  vector of ones, then we run one single replication considering sample size  $n = 50$ ; the obtained estimates from this procedure were assigned to  $\xi_0$  and used in all of the aforementioned scenarios.

The results for both parameters of the Normal-Weibull density (8), shown in Table 1, indicate that the estimates are fairly close to the actual values. Moreover, as it would be expected, the bigger the sample size, the smaller the MSEs.

**Table 1.** Bias and MSE of the estimates under the maximum likelihood method for the Normal-Weibull model.

Actual Value			Bias		MSE	
<i>n</i>	<i>k</i>	$\lambda$	$\hat{k}$	$\hat{\lambda}$	$\hat{k}$	$\hat{\lambda}$
50	1.0	1.7	0.02707850	−0.00948110	0.00822299	0.00659991
	0.5	2.0	0.01446186	−0.02037125	0.00209051	0.03647395
	3.0	0.5	0.07878343	−0.00096778	0.07352625	0.00006357
	0.9	4.0	0.02586182	−0.02752101	0.00675452	0.04497248
	7.1	5.8	0.19086399	−0.00540407	0.41402178	0.00153179
100	1.0	1.7	0.01306919	−0.00453981	0.00377883	0.00332042
	0.5	2.0	0.00726917	−0.01412373	0.00095786	0.01838681
	3.0	0.5	0.03914672	−0.00057766	0.03400929	0.00003204
	0.9	4.0	0.01167774	−0.01403219	0.00305903	0.02273089
	7.1	5.8	0.08363335	−0.00279451	0.18835856	0.00077190
200	1.0	1.7	0.00651588	−0.00253820	0.00181409	0.00166703
	0.5	2.0	0.00358578	−0.00678178	0.00045681	0.00923355
	3.0	0.5	0.01901041	−0.00029362	0.01628677	0.00001604
	0.9	4.0	0.00658567	−0.00745541	0.00148059	0.01138373
	7.1	5.8	0.03656102	−0.00066316	0.09041610	0.00038519
500	1.0	1.7	0.00317837	−0.00127234	0.00071164	0.00066754
	0.5	2.0	0.00195165	−0.00609800	0.00017967	0.00370983
	3.0	0.5	0.00748033	−0.00008804	0.00636008	0.00000641
	0.9	4.0	0.00297109	−0.00200533	0.00057744	0.00455810
	7.1	5.8	0.01427116	−0.00045889	0.03550063	0.00015444

The results given in Table 2 suggest that the estimates of the parameters of the Normal-log-logistic model (10) have similar behavior of those shown in Table 1, that is to say, the biases are quite small and the MSE decreases as the sample size increases.

**Table 2.** Bias and MSE of the estimates under the maximum likelihood method for the Normal-log-logistic model.

Actual Value			Bias		MSE	
<i>n</i>	$\alpha$	$\beta$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$
50	2.7	5.0	0.00054404	0.13424700	0.00109551	0.20695579
	0.4	1.2	0.00024567	0.03275857	0.00041864	0.01195999
	6.0	2.5	−0.00002168	0.06835325	0.02161847	0.05192641
	4.0	3.4	0.00185659	0.08997153	0.00521251	0.09539035
	1.0	8.0	0.00010181	0.21835854	0.00005868	0.53164800
100	2.7	5.0	0.00012046	0.06377031	0.00055694	0.09509012
	0.4	1.2	0.00005101	0.01519026	0.00021246	0.00547655
	6.0	2.5	0.00191769	0.03297888	0.01099844	0.02386620
	4.0	3.4	−0.00004923	0.04701171	0.00263820	0.04439236
	1.0	8.0	0.00009827	0.11057419	0.00002980	0.24582360
200	2.7	5.0	0.00015504	0.03299212	0.00028047	0.04585339
	0.4	1.2	0.00006551	0.00866507	0.00010677	0.00265802
	6.0	2.5	−0.00106277	0.01625314	0.00553891	0.01145699
	4.0	3.4	−0.00002407	0.02295567	0.00133064	0.02123514
	1.0	8.0	0.00000852	0.05090923	0.00001503	0.11715066
500	2.7	5.0	0.00021558	0.01327469	0.00011277	0.01789692
	0.4	1.2	−0.00008941	0.00435633	0.00004284	0.00104215
	6.0	2.5	−0.00042358	0.00648252	0.00222639	0.00447192
	4.0	3.4	−0.00007789	0.00855609	0.00053513	0.00826466
	1.0	8.0	0.00003954	0.01663711	0.00000604	0.04559331

### 3.2. Applications

The first data to be considered is related to the soil fertility influence and the characterization of the biologic fixation of  $N_2$  for the *Dimorphandra wilsonii* Rizz growth. It was originally studied by [24] and it also figures in the work of [25]. For 128 plants, the phosphorus concentration in the leaves was quantified. Here are the numbers: 0.22, 0.17, 0.11, 0.10, 0.15, 0.06, 0.05, 0.07, 0.12, 0.09, 0.23, 0.25, 0.23, 0.24, 0.20, 0.08, 0.11, 0.12, 0.10, 0.06, 0.20, 0.17, 0.20, 0.11, 0.16, 0.09, 0.10, 0.12, 0.12, 0.10, 0.09, 0.17, 0.19, 0.21, 0.18, 0.26, 0.19, 0.17, 0.18, 0.20, 0.24, 0.19, 0.21, 0.22, 0.17, 0.08, 0.08, 0.06, 0.09, 0.22, 0.23, 0.22, 0.19, 0.27, 0.16, 0.28, 0.11, 0.10, 0.20, 0.12, 0.15, 0.08, 0.12, 0.09, 0.14, 0.07, 0.09, 0.05, 0.06, 0.11, 0.16, 0.20, 0.25, 0.16, 0.13, 0.11, 0.11, 0.11, 0.08, 0.22, 0.11, 0.13, 0.12, 0.15, 0.12, 0.11, 0.11, 0.15, 0.10, 0.15, 0.17, 0.14, 0.12, 0.18, 0.14, 0.18, 0.13, 0.12, 0.14, 0.09, 0.10, 0.13, 0.09, 0.11, 0.11, 0.14, 0.07, 0.07, 0.19, 0.17, 0.18, 0.16, 0.19, 0.15, 0.07, 0.09, 0.17, 0.10, 0.08, 0.15, 0.21, 0.16, 0.08, 0.10, 0.06, 0.08, 0.12, 0.13. Table 3 brings some descriptive statistics.

**Table 3.** Descriptive statistics for soil fertility dataset.

n	mean	median	min	max	Variance	Skewness	Kurtosis
128	0.14078	0.13	0.05	0.28	0.00296	0.45438	−0.64478

We fitted the Normal-Weibull distribution (NW) (7) to the soil fertility dataset and compared it to the fits of Weibull (W), Exponentiated Weibull (ExpW) [1], Marshall-Olkin Extended Weibull (MOEW) [26], Kumaraswamy-Weibull (KwW) [9], Beta-Weibull (BW) [8] and McDonald-Weibull (McW) [7]. The function goodness . fit of the R package AdequacyModel provides, besides the MLEs and the standard errors (SE), some criteria for model selection (AIC, CAIC, BIC and HQIC); they are shown in Table 4.

**Table 4.** Fitted distributions to the soil fertility dataset (estimates and information criteria).

Distribution	Parameters	Estimates (SE)	AIC	CAIC	BIC	HQIC
NW	$k$	0.8398477 (0.0445182)	−395.7584	−395.6624	−390.0544	−393.4408
	$\lambda$	0.2049909 (0.0074017)				
W	$k$	2.8185566 (0.1919639)	−385.6297	−385.5337	−379.9256	−383.3121
	$\lambda$	0.1584836 (0.0052564)				
ExpW	$k$	1.5321145 (0.5023377)	−387.4361	−387.2426	−378.8801	−383.9598
	$\lambda$	0.0939938 (0.0374090)				
	$a$	3.5076974 (2.6763009)				
MOEW	$k$	3.9300962 (0.2426080)	−377.9475	−377.754	−369.3914	−374.4711
	$\lambda$	8.9163819 (4.5940983)				
	$a$	0.0031628 (0.0007240)				
KwW	$k$	1.1503912 (0.3443931)	−384.1491	−383.8239	−372.7409	−379.5139
	$\lambda$	0.1953371 (0.1291154)				
	$a$	3.3444607 (1.5352029)				
	$b$	7.5480698 (10.206142)				
BW	$k$	0.8477957 (0.2166409)	−385.7589	−385.4337	−374.3508	−381.1237
	$\lambda$	0.3304922 (0.4395169)				
	$a$	9.0436364 (4.5271059)				
	$b$	15.211970 (22.481984)				
McW	$k$	5.6665646 (8.3928707)	−384.6657	−384.1739	−370.4055	−378.8717
	$\lambda$	0.5912941 (0.5124852)				
	$a$	13.441193 (23.051917)				
	$b$	14.363802 (18.264058)				
	$c$	0.0870787 (0.1234075)				

Information criteria may be used as relative goodness-of-fit measures, such that the lowest values will characterize the best fitted models. In this sense, the Normal-Weibull distribution outperforms the other ones.

Figure 5 shows the histogram of soil fertility data and the fitted densities with the three lowest values of AIC among the distributions in the first column of Table 4. Although the Normal-Weibull and Exponentiated Weibull curves appear to be very close, the blue one (NW) seems to be closer to the histogram.

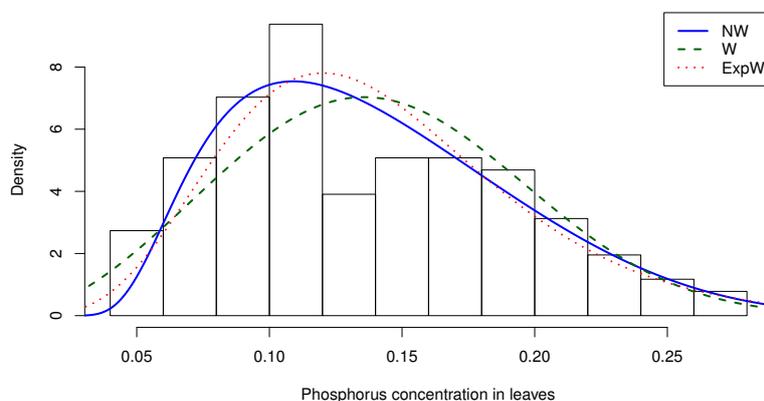


Figure 5. Histogram of soil fertility dataset and fitted densities.

The modified versions of Anderson-Darling ( $A^*$ ) and Cramér-von Mises ( $W^*$ ) statistics (more details in [27]) are typically used to investigate the quality of fit of probabilistic models. Table 5 brings these statistics concerning the fitted models to soil fertility data.

Table 5. Goodness-of-fit test statistics.

Distribution	$A^*$	$W^*$
NW	0.454008	0.079841
W	1.156994	0.207118
ExpW	0.784451	0.138403
MOEW	1.123759	0.183128
KwW	0.907239	0.163617
BW	0.750593	0.130501
McW	0.758296	0.137509

The measures portrayed in Table 5 represent the difference between the empirical distribution function and the real underlying cdf; hence we will consider that the models with lower values of  $A^*$  and  $W^*$  fit the data better. Therefore, once again the Normal-Weibull distribution beats the competing models.

The second application concerns to a dataset representing waiting times (in seconds) between 65 successive eruptions of water through a hole in the cliff at the coastal town of Kiama (New South Wales, Australia), known as the Blowhole. These data can be obtained in [17,28]. Here are they: 83, 51, 87, 60, 28, 95, 8, 27, 15, 10, 18, 16, 29, 54, 91, 8, 17, 55, 10, 35, 47, 77, 36, 17, 21, 36, 18, 40, 10, 7, 34, 27, 28, 56, 8, 25, 68, 146, 89, 18, 73, 69, 9, 37, 10, 82, 29, 8, 60, 61, 61, 18, 169, 25, 8, 26, 11, 83, 11, 42, 17, 14, 9, 12. Table 6 provides descriptive statistics.

Table 6. Descriptive statistics for eruption dataset.

n	Mean	Median	min	max	Variance	Skewness	Kurtosis
64	39.82812	28	7	169	1139.097	1.54641	2.77108

We fitted the Normal-log-logistic distribution (NLL) (9) to the eruption dataset and compared it to the fits of Log-logistic (LL), Exponentiated Log-logistic (ExpLL), Beta-log-logistic (BLL),

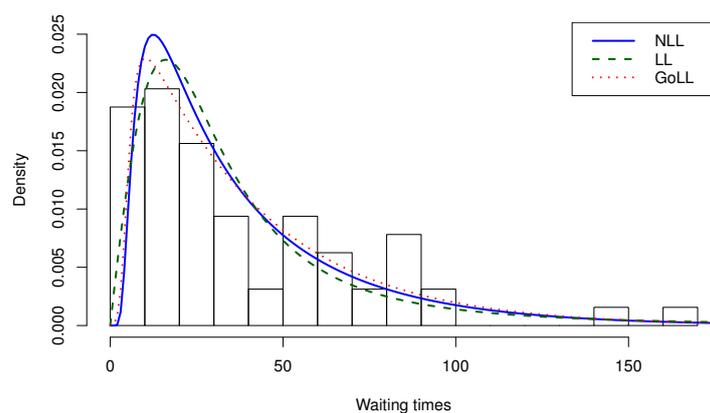
Kumaraswamy-log-logistic (KwLL) and Gompertz-log-logistic (GoLL); the four latter along the lines of [1,8,9,11] respectively. Table 7 brings the MLEs, SEs and information criteria.

**Table 7.** Fitted distributions to the eruption dataset (estimates and information criteria).

Distribution	Parameters	Estimates (SE)	AIC	CAIC	BIC	HQIC
NLL	$\alpha$	28.71747 (2.751091)	587.5681	587.7649	591.8859	589.2691
	$\beta$	0.568200 (0.042998)				
LL	$\alpha$	28.27831 (3.203986)	597.1497	597.3464	601.4674	598.8506
	$\beta$	1.969345 (0.198878)				
ExpLL	$\alpha$	7.394859 (6.479904)	597.3629	597.7629	603.8396	599.9144
	$\beta$	1.461528 (0.197256)				
	$a$	4.572569 (4.470086)				
BLL	$\alpha$	7.445103 (10.72863)	596.186	596.864	604.8215	599.588
	$\beta$	0.484528 (0.223626)				
	$a$	17.28664 (13.82502)				
	$b$	9.285354 (9.566756)				
KwLL	$\alpha$	2.107772 (5.325557)	596.68	597.358	605.3156	600.082
	$\beta$	0.511324 (0.130629)				
	$a$	12.14489 (12.25210)				
	$b$	11.42477 (8.749787)				
GoLL	$\alpha$	5.667167 (1.680598)	591.7172	592.3952	600.3528	595.1192
	$\beta$	4.348435 (1.450980)				
	$a$	0.035617 (0.017111)				
	$b$	0.234894 (0.100666)				

Since the Normal-log-logistic fitted model presents the smallest values of AIC, CAIC, BIC and HQIC compared to the fits of the other distributions, selecting it rather than the others is a reasonable decision in this case.

In Figure 6 the histogram of eruption data and the fitted densities with the three lowest values of AIC among the distributions in the first column of Table 7 are depicted. By a visual comparison, the three curves are apparently good approximations to the histogram, but the Normal-log-logistic's seems to explain the behavior of the data more accurately.



**Figure 6.** Histogram of eruption dataset and fitted densities.

Table 8 provides the values of  $A^*$  and  $W^*$  of the distributions in the first column of Table 7. These statistics suggest that GoLL and NLL models fit the eruption dataset very closely. Nonetheless, in order to pick a more parsimonious model, one should prefer the NLL, since it has fewer parameters than GoLL.

**Table 8.** Goodness-of-fit tests.

Distribution	A*	W*
NLL	0.612291	0.0803799
LL	1.019129	0.1413872
ExpLL	1.138218	0.1617136
BLL	0.837211	0.1141264
KwLL	0.818072	0.1118931
GoLL	0.605822	0.0805111

It is worth mentioning that [17] proposed the new class Exponentiated Kumaraswamy-G and fitted one of its submodels (with Weibull as baseline) to the same eruption dataset. It presented  $A^* = 0.7594$  and  $W^* = 0.1037$ , whereas NLL presented lower values of these statistics as one can check in Table 8.

#### 4. Concluding Remarks

Based on the method of generating classes of probability distributions presented by [12], we introduce a new class called Normal-G. It has the advantage of demanding no additional parameters besides the baseline ones. We demonstrate that the proposed class generates identifiable sub-models as long as the parent distribution is identifiable. The pdf of the class can be written as a linear combination of pdfs of exponentiated distributions; it allows us to easily derive the raw moments, the incomplete moments and the moment generating function.

We bring Monte Carlo simulation studies to attest the good performance of the MLEs of two distributions generated by the class and to illustrate its usefulness, applications to real datasets are made. The fitted models are compared to other competitive distributions regarding the Anderson-Darling and the Cramér-von Mises statistics, as well as commonly used information criteria as goodness-of-fit measures. The general results indicate that the Normal-G outperforms the other distributions in comparison. The new class is powerful and provides parsimonious models, which may hopefully interest practitioners of statistics, soil science, oceanography and other fields.

**Author Contributions:** All of the authors contributed relevantly to this research article.

**Funding:** This research received no external funding.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A

The information matrix mentioned in Section 2.5 is given by  $J(\xi|\mathbf{X}) = -\nabla_{\xi} \nabla_{\xi}^{\top} \ell(\xi|\mathbf{X}) = -(u_{ih})_{1 \leq i \leq r, 1 \leq h \leq r}$ , where:

$$\begin{aligned} u_{ih} = & \sum_{j=1}^m \left[ \frac{2G_{\xi}(x_j) - 3}{G_{\xi}^4(x_j)} - \frac{2G_{\xi}(x_j) + 1}{[1 - G_{\xi}(x_j)]^4} \right] \cdot \frac{\partial}{\partial \xi_i} G_{\xi}(x_j) \cdot \frac{\partial}{\partial \xi_h} G_{\xi}(x_j) \\ & + \sum_{j=1}^m \frac{[1 - G_{\xi}(x_j)]^4 - G_{\xi}^4(x_j)}{G_{\xi}^3(x_j)[1 - G_{\xi}(x_j)]^3} \cdot \frac{\partial^2}{\partial \xi_i \partial \xi_h} G_{\xi}(x_j) \\ & + \sum_{j=1}^m \frac{8[1 - G_{\xi}(x_j)]G_{\xi}(x_j)}{1 - 2G_{\xi}(x_j) + 2G_{\xi}^2(x_j)} \cdot \frac{\partial}{\partial \xi_i} G_{\xi}(x_j) \cdot \frac{\partial}{\partial \xi_h} G_{\xi}(x_j) \\ & + \sum_{j=1}^m \frac{4G_{\xi}(x_j) - 2}{1 - 2G_{\xi}(x_j) + 2G_{\xi}^2(x_j)} \cdot \frac{\partial^2}{\partial \xi_i \partial \xi_h} G_{\xi}(x_j) \\ & + \sum_{j=1}^m \frac{2}{G_{\xi}^2(x_j)} \cdot \frac{\partial}{\partial \xi_i} G_{\xi}(x_j) \cdot \frac{\partial}{\partial \xi_h} G_{\xi}(x_j) - \sum_{j=1}^m \frac{2}{G_{\xi}(x_j)} \cdot \frac{\partial^2}{\partial \xi_i \partial \xi_h} G_{\xi}(x_j) \\ & + \sum_{j=1}^m \frac{2}{[1 - G_{\xi}(x_j)]^2} \cdot \frac{\partial}{\partial \xi_i} G_{\xi}(x_j) \cdot \frac{\partial}{\partial \xi_h} G_{\xi}(x_j) + \sum_{j=1}^m \frac{2}{1 - G_{\xi}(x_j)} \cdot \frac{\partial^2}{\partial \xi_i \partial \xi_h} G_{\xi}(x_j) \\ & - \sum_{j=1}^m \frac{1}{g_{\xi}^2(x_j)} \cdot \frac{\partial}{\partial \xi_i} g_{\xi}(x_j) \cdot \frac{\partial}{\partial \xi_h} g_{\xi}(x_j) + \sum_{j=1}^m \frac{1}{g_{\xi}(x_j)} \cdot \frac{\partial^2}{\partial \xi_i \partial \xi_h} g_{\xi}(x_j). \end{aligned}$$

## References

- Mudholkar, G.S.; Srivastava, D.K.; Freimer M. The exponentiated Weibull family: A reanalysis of the bus motor failure data. *Technometrics* **1995**, *37*, 436–445. [\[CrossRef\]](#)
- Gupta, R.D.; Kundu D. Generalized Exponential Distributions. *Aust. N. Z. J. Stat.* **1999**, *41*, 173–188. [\[CrossRef\]](#)
- Marshall, A.W.; Olkin I. A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families. *Biometrika* **1997**, *84*, 641–652. [\[CrossRef\]](#)
- Nadarajah, S. A Generalized Normal Distribution. *J. Appl. Stat.* **2005**, *32*, 685–694. [\[CrossRef\]](#)
- Azzalini, A. A Class of Distributions which includes the Normal ones. *Scand. J. Stat.* **1985**, *12*, 171–178.
- Robertson, H.T.; Allison, D.B. A Novel Generalized Normal Distribution for Human Longevity and other Negatively Skewed Data. *PLoS ONE* **2012**, *7*, e37025. [\[CrossRef\]](#)
- Cordeiro, G.M.; Hashimoto, E.M.; Ortega, E.M.M. The McDonald Weibull Model. *Statistics* **2014**, *48*, 256–278. [\[CrossRef\]](#)
- Famoye, F.; Lee, C.; Olumolade, O. The Beta-Weibull distribution. *J. Stat. Theory Appl.* **2005**, *4*, 121–136.
- Cordeiro, G.M.; Castro, M. A new family of generalized distributions. *J. Stat. Comput. Simul.* **2010**, *81*, 883–898. [\[CrossRef\]](#)
- Alzaatreh, A.; Lee, C.; Famoye, F. A new method for generating families of continuous distributions. *Metron* **2013**, *71*, 63–79. [\[CrossRef\]](#)
- Alizadeh, M.; Cordeiro, G.M.; Pinho, L.G.B.; Ghosh, I. The Gompertz-G family of distributions. *J. Stat. Theory Pract.* **2017**, *11*, 179–207. [\[CrossRef\]](#)
- Brito, C.R.; Rêgo, L.C.; Oliveira, W.R.; Gomes-Silva F. Method for Generating Distributions and Classes of Probability Distributions: The Univariate Case. *Hacet. J. Math. Stat.* **2019**, *48*, 897–930.
- Xie, M.; Tang, Y.; Goh T.N. A modified Weibull extension with bathtub-shaped failure rate function. *Reliab. Eng. Syst. Safe.* **2002**, *76*, 279–285. [\[CrossRef\]](#)
- Bebbington, M.; Lai, C.D.; Zitikis, R. A flexible Weibull extension. *Reliab. Eng. Syst. Safe.* **2007**, *92*, 719–726. [\[CrossRef\]](#)
- Tahir, M.H.; Nadarajah, S. Parameter induction in continuous univariate distributions: Well-established G families. *An. Acad. Bras. Ciênc.* **2015**, *87*, 539–568. [\[CrossRef\]](#)

16. Cordeiro, G.M.; Ortega, E.M.M.; Cunha, D.C.C. The Exponentiated Generalized Class of Distributions. *J. Data Sci.* **2013**, *11*, 1–27.
17. Silva, R.; Gomes-Silva F.; Ramos, M.; Cordeiro, G.; Marinho, P.; Andrade, T.A.N. The Exponentiated Kumaraswamy-G Class: General Properties and Application. *Rev. Colomb. Estad.* **2019**, *42*, 1–33. [[CrossRef](#)]
18. Cakmakyapan, S.; Ozel, G. The Lindley Family of Distributions: Properties and Applications. *Hacet. J. Math. Stat.* **2017**, *46*, 1113–1137. [[CrossRef](#)]
19. Barreto-Souza, W.; Cordeiro, G.M.; Simas, A.B. Some Results for Beta Fréchet Distribution. *Commun. Stat. Theory Methods* **2011**, *40*, 798–811. [[CrossRef](#)]
20. Huang, S.; Oluyede, B.O. Exponentiated Kumaraswamy-Dagum distribution with applications to income and lifetime data. *J. Stat. Dist. Appl.* **2014**, *1*, 1–20. [[CrossRef](#)]
21. Cordeiro, G.M.; Bager, R.S.B. Moments for Some Kumaraswamy Generalized Distributions. *Comm. Statist. Theory Methods* **2015** *44*, 2720–2737. [[CrossRef](#)]
22. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2018.
23. Von Neumann, J. Various techniques used in connection with random digits. In *Applied Mathematics Series 12*; National Bureau of Standards: Washington, DC, USA, 1951; pp. 36–38.
24. Fonseca, M.B. A influência da fertilidade do solo e caracterização da fixação biológica de N<sub>2</sub> para o crescimento de *Dimorphandra wilsonii* Rizz. Master's Thesis, Federal University of Minas Gerais, Belo Horizonte, Brazil, 2007.
25. Silva, R.B.; Bourguignon, M.; Dias, C.R.B.; Cordeiro, G.M. The compound class of extended Weibull power series distributions. *Comput. Stat. Data Anal.* **2013**, *58*, 352–367. [[CrossRef](#)]
26. Cordeiro, G.M.; Lemonte, A.J. On the Marshall–Olkin extended Weibull distribution. *Stat. Pap.* **2013**, *54*, 333–353. [[CrossRef](#)]
27. Chen, G.; Balakrishnan, N. A general purpose approximate goodness-of-fit test. *J. Qual. Technol.* **1995**, *27*, 154–161. [[CrossRef](#)]
28. StatSci.org. Available online: <http://www.statsci.org/data/oz/kiama.html> (accessed on 21 October 2019).



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).