

Review

An In-Depth Review of the Weibull Model with a Focus on Various Parameterizations

Yolanda M. Gómez ^{1,†} , Diego I. Gallardo ^{1,*,†} , Carolina Marchant ^{2,†} , Luis Sánchez ^{3,†} 
and Marcelo Bourguignon ^{4,†} 

¹ Departamento de Estadística, Facultad de Ciencias, Universidad del Bío-Bío, Concepción 4081112, Chile; ygomez@ubiobio.cl

² Faculty of Basic Sciences, Universidad Católica del Maule, Talca 3480112, Chile; cmarchant@ucm.cl

³ Institute of Statistics, Universidad Austral de Chile, Valdivia 5110566, Chile; luis.sanchez@uach.cl

⁴ Statistics Department, Federal University of Rio Grande do Norte, Natal 59078-970, RN, Brazil; marcelo.bourguignon@ufrn.br

* Correspondence: dgallardo@ubiobio.cl

† These authors contributed equally to this work.

Abstract: The Weibull distribution is a versatile probability distribution widely applied in modeling the failure times of objects or systems. Its behavior is shaped by two essential parameters: the shape parameter and the scale parameter. By manipulating these parameters, the Weibull distribution adeptly captures diverse failure patterns observed in real-world scenarios. This flexibility and broad applicability make it an indispensable tool in reliability analysis and survival modeling. This manuscript explores five parameterizations of the Weibull distribution, each based on different moments, like mean, quantile, and mode. It meticulously characterizes each parameterization, introducing a novel one based on the model's mode, along with its hazard and survival functions, shedding light on their unique properties. Additionally, it delves into the interpretation of regression coefficients when incorporating regression structures into these parameterizations. It is analytically established that all five parameterizations define the same log-likelihood function, underlining their equivalence. Through Monte Carlo simulation studies, the performances of these parameterizations are evaluated in terms of parameter estimations and residuals. The models are further applied to real-world data, illustrating their effectiveness in analyzing material fatigue life and survival data. In summary, this manuscript provides a comprehensive exploration of the Weibull distribution and its various parameterizations. It offers valuable insights into their applications and implications in modeling failure times, with potential contributions to diverse fields requiring reliability and survival analysis.

Keywords: regression models; reparameterization; survival modeling; Weibull distribution

MSC: 62N02; 62J99



Citation: Gómez, Y.M.; Gallardo, D.I.; Marchant, C.; Sánchez, L.; Bourguignon, M. An In-Depth Review of the Weibull Model with a Focus on Various Parameterizations.

Mathematics **2024**, *12*, 56. <https://doi.org/10.3390/math12010056>

Academic Editor: Antonio Di Crescenzo

Received: 8 November 2023

Revised: 12 December 2023

Accepted: 21 December 2023

Published: 23 December 2023



Copyright: © 2023 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 (<https://creativecommons.org/licenses/by/4.0/>).

1. Background

The Weibull distribution is a popular continuous probability distribution that is commonly utilized to model the lifetimes or failure times of objects or systems. It was initially introduced by Waloddi Weibull in 1951 [1] and has since found applications in various fields. For instance, in reliability engineering, Keshevan et al. [2] employed the Weibull distribution to model the Hertzian fracture of Pyrex glass. In another study, Queeshi and Sheikh [3] used the Weibull distribution to analyze adhesive wear in metals. Similarly, Durham and Padgett [4] applied the distribution to carbon fibers and small composite specimens. Almeida [5] investigated the failure of coating using the Weibull distribution, while Fok et al. [6] focused on its use for brittle material. Additionally, Newell et al. [7]

employed the distribution to study compressive failure in high-performance polymers and Li et al. [8] analyzed concrete components using the Weibull distribution.

The distribution mentioned above has been widely applied as a flexible modeling tool in various fields, addressing a diverse range of issues. It has been successfully utilized in disciplines such as quality control, weather forecasting, industrial engineering, electric systems engineering, communications systems engineering, hydrology, and more. For instance, Bebbington and Lai [9] utilized it for volcanic eruptions, while Durrans [10] and Heo et al. [11] applied it to regional flood frequency analysis. Fleming [12] used it to describe the dynamics of foliage biomass on Scots pine. In the field of Economics, Roed and Zhang [13] employed it to analyze unemployment duration data. In the context of wireless communications, the Weibull distribution is very flexible. Ikki and Ahmed [14] analyzed the performance of multi-hop relaying systems over Weibull fading channels in terms of bit error rate and outage probability. A similar analysis of the bit error rate and outage probability over multi-hop Weibull fading channels was also conducted by Wang et al. [15].

The Weibull distribution is defined by two parameters: the shape parameter (ν) and the scale parameter (λ). The shape parameter determines the shape of the distribution curve, while the scale parameter determines the characteristic magnitude or scale of the failure times. Depending on the value of the shape parameter ν , the Weibull distribution can exhibit different shapes. When $\nu > 1$, the distribution is positively skewed, indicating a decreasing failure rate over time. This shape is commonly known as the bathtub curve and is often observed in reliability analysis. In this curve, failures are more likely to occur either early on, due to manufacturing defects or initial wear, or in the later stages, due to aging or wear-out effects. When $\nu = 1$, the distribution simplifies to the exponential distribution, which is the only continuous distribution with a constant hazard function on the positive axis. For $\nu < 1$, the distribution is negatively skewed, indicating an increasing failure rate over time (Rinne [16]).

The versatility and importance of this distribution can be seen in its close relationship with several well-known distributions in statistics. It includes other distributions as special cases, such as the exponential distribution (when the shape parameter ν is equal to 1) and the Rayleigh distribution (when ν is equal to 2). Furthermore, if T follows a Weibull distribution, then $Y = \lambda(1 - \nu \log(T/\nu))$ follows an extreme-value distribution. By applying a simple log transformation, the Weibull distribution can also be converted into the Gumbel distribution. Additionally, it acts as a limit distribution for the Burr distribution, establishing a significant connection between these distributions (Lai and Xie [17]). These relationships further emphasize the importance and wide-ranging applicability of the Weibull distribution in statistical modeling and analysis.

The Weibull distribution is widely used in modeling lifetimes because it can effectively represent different failure patterns observed in real-life situations. Failure data can be classified into two types: complete data and incomplete (censored) data. Complete data refer to cases where the actual observed values are known for every observation in the dataset. On the other hand, censored data occur when the actual observed values are unknown for some or all of the observations. It is worth noting that there are various types of censoring; more information can be found in the work by Lawless [18]. Several studies have examined the application of the Weibull distribution in analyzing censored data. For example, Ghitany et al. (2005) [19] and Klakattawi (2022) [20] extended the Weibull model and applied it to bladder cancer data, Joarder (2011) [21] and Jia et al. (2016) [22] studied leukemia data, and Lee et al. (2007) [23] focused on the head-and-neck-cancer trial. The Weibull distribution has an intriguing application in models involving a fraction of cure. This distribution is particularly well-suited for estimating the time it takes for cancer cells to produce detectable cancer. As a result, it has gained significant popularity in this field. Several studies, including those by Chen et al. [24], Yin and Ibrahim [25], Rodrigues et al. [26], Gallardo et al. [27], and Azimi et al. [28], have extensively explored this topic.

There are numerous extensions, generalizations, and modifications to the Weibull distribution. These developments have emerged to address the requirements of empirical datasets that exhibit characteristics beyond what can be effectively captured by a standard two-parameter Weibull model (Lai and Xie [17]). These extended models can be broadly classified into three groups: univariate, multivariate, and stochastic process models. Univariate models focus on enhancing the flexibility of the Weibull distribution for single variable analysis. Multivariate models extend the Weibull framework to handle multiple variables and their dependencies. Stochastic process models delve into time-dependent and dynamic variations of the Weibull distribution. Noteworthy references, such as Murthy et al. [29,30], provide insights into these diverse extensions and shed light on the advancements made in modeling techniques beyond the traditional Weibull framework. Silva et al. [31] derived the power-series extended Weibull class of distributions. Santos-Neto et al. [32] and Nascimento et al. [33] introduced a family of distributions encompassing well over forty variants. For more details about the generalizations and modifications of the Weibull distribution, see Murthy [30], Nadarajah and Kotz [34], Pham and Lai [35], and Almalki and Nadarajah [36], as well as the references therein. The analysis of the truncated Weibull distribution has been explored in many papers. For example, Wingo [37] proposed the left-truncated Weibull distribution. The right-truncated Weibull distribution has been analyzed by Zhang and Xie [38]. The doubly-truncated Weibull distribution has been studied in work by McEwen and Parresol [39].

In summary, the Weibull distribution is a versatile probability distribution that finds widespread application in modeling the failure times of objects or systems. The distribution's characteristics and behaviors are determined by two key parameters: the shape parameter and the scale parameter. The shape parameter governs the shape of the distribution, allowing it to range from exponential to highly skewed distributions. On the other hand, the scale parameter influences the location and spread of the distribution. By manipulating these parameters, the Weibull distribution can effectively capture various failure patterns observed in real-world scenarios. Its flexibility and wide range of applications make it a valuable tool in reliability analysis and survival modeling.

The main objectives of this manuscript are as follows:

1. We perform a review of the different parameterizations of the Weibull distribution documented in the literature, including the interpretation of the regression coefficients, when incorporating regression structures into these parameterizations.
2. We introduce a novel parameterization of the Weibull model based on the mode of this distribution.
3. We theoretically explore the equivalence of the five parameterizations of the Weibull distribution in the context of regression models, since there is no discussion connecting all the mentioned parameterizations.

The manuscript is organized as follows. In Section 2, five parameterizations of the Weibull distribution are introduced. Three of these parameterizations are mean-based, one is quantile-based, and the last one is mode-based. Each parameterization is characterized by examining the respective hazard and survival functions. Section 3 delves into the interpretation of regression coefficients when regression structures are incorporated into the parameters of the five Weibull model parameterizations. In Section 5, we analytically demonstrate that the five parameterizations of the Weibull model discussed in the previous section define the same log-likelihood function. Section 6 presents Monte Carlo (MC) simulation studies for parameter estimation and residuals. In Section 6, the analyzed models are applied to real data. Specifically, two illustrations are provided to exemplify their applicability and use in material fatigue life and survival data. It should be noted that all models have been implemented using the R-project code (<https://www.r-project.org/>, accessed on 5 December 2023). Finally, Section 7 summarizes the main conclusions of the manuscript and discusses possible avenues for future work.

2. Weibull Parameterizations

As discussed in Section 1, the Weibull distribution is one of the most used models in reliability analysis because it is a parsimonious model with a simple expression for the density, survival, and hazard functions, and with some interesting properties. For instance, it can assume a decreasing, increasing, or constant hazard rate, depending only on the shape parameter ν (<1 , >1 , or $=1$, respectively). However, the Weibull distribution has many parameterizations in the literature, depending on the study.

The accelerated failure time (AFT) model, which is employed as the base parameterization, is used for the Weibull model; it has hazard and survival functions given by

$$h(t; \lambda, \nu) = \frac{\nu}{\lambda} \left[\frac{t}{\lambda} \right]^{\nu-1}, \quad t, \nu, \lambda > 0.$$

and

$$S(t; \lambda, \nu) = \exp\left(-\left[\frac{t}{\lambda}\right]^\nu\right), \quad t, \nu, \lambda > 0,$$

respectively, with λ and ν parameters of scale and shape, respectively. This parameterization is referred to as WEI(λ, ν). An alternative parameterization of the Weibull model is associated with the proportional hazards model, in which the hazard and survival functions are given by

$$h(t; \lambda, \nu) = \nu \lambda t^{\nu-1}, \quad t, \nu, \lambda > 0$$

and

$$S(t; \lambda, \nu) = \exp(-\lambda t^\nu), \quad t, \nu, \lambda > 0,$$

respectively. This parameterization is referred to as WEI2(λ, ν), where λ and ν act as the scale and shape parameters, respectively. A third alternative parameterization of the Weibull model is related to its mean. Fernandes et al. [40] propose a reparameterization that expresses the Weibull distribution in terms of the process mean, enabling straightforward monitoring of the Weibull mean. In this parameterization, the hazard and survival functions are given by

$$h(t; \lambda, \nu) = \frac{\nu}{\gamma} \left[\frac{t}{\gamma} \right]^{\nu-1}, \quad t, \nu, \lambda > 0$$

where $\gamma = \lambda/\Gamma(1 + 1/\nu)$ and

$$S(t; \lambda, \nu) = \exp\left(-\left[\frac{t}{\gamma}\right]^\nu\right), \quad t, \nu, \lambda > 0,$$

respectively. This parameterization is referred to as WEI3(λ, ν). In this case, $\mathbb{E}(T) = \lambda$.

Recently, a fourth parameterization for the Weibull distribution was introduced by Sánchez et al. [41]. In this parameterization, the hazard and survival functions are given by

$$h(t; \lambda, \nu) = -\log(1 - q) \frac{\nu}{\lambda} \left[\frac{t}{\lambda} \right]^{\nu-1}, \quad t, \nu, \lambda > 0.$$

where $q \in (0, 1)$ and

$$S(t; \lambda, \nu) = \exp\left(\log(1 - q) \left[\frac{t}{\lambda}\right]^\nu\right), \quad t, \nu, \lambda > 0,$$

respectively. This parameterization is referred to as WEI4(λ, ν). In this case, λ represents the q -quantile of the distribution.

Alternatively, a fifth parameterization of the Weibull is proposed, based on its mode. The main motivation for re-parameterizing the model in terms of this measure is associated with the robustness of the mode (above the mean, for example) and the fact that it is a measure that has been used more frequently in the literature in recent years. For instance,

see Yao and Li [42], Chen [43], and Bourguignon et al. [44]. In this parameterization, the hazard and survival functions are given by

$$h(t; \lambda, \nu) = \frac{[\nu - 1]}{\lambda} \left[\frac{t}{\lambda} \right]^{\nu-1}, \quad t, \nu, \lambda > 0.$$

and

$$S(t; \lambda, \nu) = \exp\left(-\left[1 - \frac{1}{\nu}\right] \left[\frac{t}{\lambda}\right]^\nu\right), \quad t, \lambda > 0, \nu > 1.$$

This parameterization is referred to as WEI5(λ, ν). In this case, λ represents the mode of the distribution. Table 1 presents the mean, mode, q -quantile, and variance for the five parameterizations of the Weibull distribution explored in this study.

Table 1. Different measures for the five parameterization of the Weibull distribution. In all parameterizations, the parameter space for (λ, ν) is \mathbb{R}_+^2 , except for the WEI5 model, where the parameter space is $\mathbb{R}_+ \times (1, \infty)$.

Model	Mean	Mode	q -Quantile	Variance
WEI	$\lambda \Gamma\left(1 + \frac{1}{\nu}\right)$	$\begin{cases} \lambda \left[\frac{(\nu-1)}{\nu}\right]^{1/\nu}, & \text{if } \nu > 1 \\ 0, & \text{if } \nu \leq 1 \end{cases}$	$\lambda [-\log(1 - q)]^{1/\nu}$	$\lambda^2 \omega(\nu)$
WEI2	$\lambda^{-1/\nu} \Gamma\left(1 + \frac{1}{\nu}\right)$	$\begin{cases} \left[\frac{(\nu-1)}{\lambda \nu}\right]^{1/\nu}, & \text{if } \nu > 1 \\ 0, & \text{if } \nu \leq 1 \end{cases}$	$\left[\frac{-\log(1-q)}{\lambda}\right]^{1/\nu}$	$\lambda^{-2/\nu} \omega(\nu)$
WEI3	λ	$\begin{cases} \frac{\lambda}{\Gamma\left(1 + \frac{1}{\nu}\right)} \left[\frac{(\nu-1)}{\nu}\right]^{1/\nu}, & \text{if } \nu > 1 \\ 0, & \text{if } \nu \leq 1 \end{cases}$	$\frac{\lambda}{\Gamma\left(1 + \frac{1}{\nu}\right)} [-\log(1 - q)]^{1/\nu}$	$\lambda^2 \frac{\omega(\nu)}{\Gamma^2\left(1 + \frac{1}{\nu}\right)}$
WEI4	$\frac{\lambda \Gamma\left(1 + \frac{1}{\nu}\right)}{[-\log(1 - q)]^{1/\nu}}$	$\begin{cases} \lambda \left[\frac{(\nu-1)}{\nu [-\log(1 - q)]}\right]^{1/\nu}, & \text{if } \nu > 1 \\ 0, & \text{if } \nu \leq 1 \end{cases}$	λ	$\lambda^2 [\log^2(1 - q)]^{-2/\nu} \omega(\nu)$
WEI5	$\lambda \left[1 - \frac{1}{\nu}\right]^{-1/\nu} \Gamma\left(1 + \frac{1}{\nu}\right)$	λ	$\lambda \left[\frac{-\nu \log(1 - q)}{(\nu - 1)}\right]^{1/\nu}$	$\lambda^2 \left[1 - \frac{1}{\nu}\right]^{-1/\nu} \omega(\nu)$

NOTE: $\omega(\nu) = \Gamma\left(1 + \frac{2}{\nu}\right) - \Gamma^2\left(1 + \frac{1}{\nu}\right)$.

For a set of p covariates observed for each individual and the intercept term, say $\mathbf{x}_i^\top = (1, x_{1i}, x_{2i}, \dots, x_{pi})$, and for the four parameterizations, it is typical to include a regression structure in λ , as follows

$$\lambda_i = g(\mathbf{x}_i^\top \boldsymbol{\beta}), \quad i = 1, \dots, n, \tag{1}$$

where $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \dots, \beta_p)$, and $g(\cdot)$ is a twice differentiable function, such as $g : \mathbb{R} \rightarrow \mathbb{R}^+$. The most common choice is $g(u) = \exp(u)$, which also facilitates the interpretation of the regression coefficients. In each model, such interpretation depends on its specific parameterization (WEI, WEI2, WEI3, WEI4, and WEI5). However, it will be demonstrated that when a regression structure is incorporated solely into λ , as in Equation (1), without affecting the shape parameter ν , then all five models share the same probability density function (PDF) and, consequently, the same log-likelihood function. Furthermore, it will also be demonstrated that introducing a dual regression structure for both λ and ν in the WEI, WEI2, WEI3, and WEI4 models results in different PDFs among the models.

3. Interpretation of the Coefficients

In this section, we explore the varying interpretations of regression coefficients in the Weibull model across different parameterizations. Let us consider two individuals with associated covariates, $\mathbf{x}_i = (1, x_{i1}, \dots, x_{ij}, \dots, x_{ip})$ and $\mathbf{x}_i^* = (1, x_{i1}, \dots, x_{ij} + 1, \dots, x_{ip})$, which is that \mathbf{x}_i and \mathbf{x}_i^* are identical, except for an increase of one unit in the j -th covariate. The scenario where the j -th covariate is quantitative will be focused on. However, in all cases, a similar interpretation can be applied to x_{ij} , assuming two categories (labeled as 0 and 1). Note that $\mathbf{x}_i^{*\top} \boldsymbol{\beta} = \mathbf{x}_i^\top \boldsymbol{\beta} + \beta_j$.

3.1. WEI Model

For the AFT parameterization, we obtain

$$S(t; \lambda_i, \nu) = \exp\left(-\left[\frac{t}{\lambda_i}\right]^\nu\right) = S_0(\lambda_i^{-1}t; \nu), \quad t > 0,$$

where $S_0(t) = \exp(-t^\nu)$. In this context, λ_i^{-1} is known as the accelerator factor. More precisely, as mentioned by Kleinbaum and Klein [45], “the accelerator factor is a ratio of survival times corresponding to any fixed value of $S(t)$ ”. In this case, to facilitate the interpretation of the coefficients, it is set $g(u) = \exp(u)$. Consequently, the interpretations of covariates are as follows: If $\exp(\beta_j) = 2$, then the median survival time doubles when the j -th covariate is increased by one unit, compared to the median survival time when the j -th covariate remains unchanged.

Remark 1. For WEI2, WEI3, and WEI4 parameterizations, the traditional link function is $g(u) = \exp(u)$.

3.2. WEI2 Model

In this case, the quotient between the hazard function related to \mathbf{x}_i and \mathbf{x}_i^* is

$$\frac{h(t; \mathbf{x}_i, \nu)}{h(t; \mathbf{x}_i^*, \nu)} = \frac{\exp(\mathbf{x}_i^{*\top} \boldsymbol{\beta}) \nu t^{\nu-1}}{\exp(\mathbf{x}_i^\top \boldsymbol{\beta}) \nu t^{\nu-1}} = \exp(\beta_j).$$

This does not depend on t (hence, the name ‘proportional hazard models’). Therefore, $\exp(\beta_j)$ should be interpreted as the increase (or decrease) in the hazard function when the j -th covariate increases by one unit.

3.3. WEI3 Model

In this case, the expected value of the distribution is $\mathbb{E}(T_i; \mathbf{x}_i) = \lambda_i$. Therefore, the quotient between the mean related to \mathbf{x}_i and \mathbf{x}_i^* is

$$\frac{\mathbb{E}(T_i; \mathbf{x}_i^*)}{\mathbb{E}(T_i; \mathbf{x}_i)} = \frac{\exp(\mathbf{x}_i^{*\top} \boldsymbol{\beta})}{\exp(\mathbf{x}_i^\top \boldsymbol{\beta})} = \exp(\beta_j).$$

Therefore, $\exp(\beta_j)$ should be interpreted as the percentage increase (or decrease) in the mean when the j -th covariate is increased by one unit.

3.4. WEI4 Model

In this case, the $100 \times q$ -th quantile of the distribution is $\tau_q(\mathbf{x}_i) = \lambda_i$. Therefore, the quotient between the $100 \times q$ -th quantile related to \mathbf{x}_i and \mathbf{x}_i^* is given by

$$\frac{\tau_q(\mathbf{x}_i^*)}{\tau_q(\mathbf{x}_i)} = \frac{\exp(\mathbf{x}_i^{*\top} \boldsymbol{\beta})}{\exp(\mathbf{x}_i^\top \boldsymbol{\beta})} = \exp(\beta_j).$$

Then, $\exp(\beta_j)$ represents the percentage increase (or decrease) in the $100 \times q$ -th quantile when the j -th covariate is increased by one unit.

3.5. WEI5 Model

For this model, the mode of the distribution is $\text{Mo}(T_i; \mathbf{x}_i) = \lambda_i$. Therefore, the model related to \mathbf{x}_i and \mathbf{x}_i^* is

$$\frac{\text{Mo}(T_i; \mathbf{x}_i^*)}{\text{Mo}(T_i; \mathbf{x}_i)} = \frac{\exp(\mathbf{x}_i^{*\top} \boldsymbol{\beta})}{\exp(\mathbf{x}_i^\top \boldsymbol{\beta})} = \exp(\beta_j).$$

In this case, $\exp(\beta_j)$ represents the percentage increase (or decrease) in the mode when the j -th covariate is increased by one unit.

3.6. The Bi-Univocity among the Four Parameterizations When Modeling Only λ

To date, there is no discussion in the literature about the connection between the different parameterizations of the Weibull distribution when incorporating regression structures into these parameterizations. The following theorem establishes the biunivocal relationship between the WEI, WEI2, WEI3, WEI4, and WEI5 models when modeling only λ .

Theorem 1. Let Y_1, \dots, Y_n be independent random variables, such as $Y_i \sim \text{WEI}(\lambda_i, \nu)$, where $\lambda_i = \exp(-\mathbf{x}_i^\top \boldsymbol{\beta})$, $i = 1, \dots, n$ and a constant shape parameter ν . Consider alternative parameterizations for the Weibull model, such as $Y_i \sim \text{WEI2}(\lambda_i^*, \tau)$, $Y_i \sim \text{WEI3}(\lambda_i^*, \tau)$, $Y_i \sim \text{WEI4}(\lambda_i^*, \tau)$ or $Y_i \sim \text{WEI5}(\lambda_i^*, \tau)$, with $\lambda_i^* = \exp(\mathbf{x}_i^\top \boldsymbol{\delta})$ and a non-modeled τ . If intercept terms are considered in \mathbf{x}_i , then the elements of $(\boldsymbol{\delta}^\top, \tau)$ can be obtained uniquely from $(\boldsymbol{\beta}^\top, \nu)$.

Proof. Equating the hazard functions for the WEI2, WEI3, WEI4, and WEI5 models with those associated with the WEI model, we obtain $\lambda_i = k(\lambda_i^*, \tau)$ and $\nu = \tau$, where $k(\lambda_i^*, \tau) = [\lambda_i^*]^{-1/\tau}$, for the WEI2 model, $k(\lambda_i^*, \tau) = \lambda_i^* \times [\tau \{\Gamma(1 + 1/\tau)\}^\tau]^{-1}$, for the WEI3 model, $k(\lambda_i^*, \tau) = \lambda_i^* \times -\log(1 - q)$, for the WEI4 model and $k(\lambda_i^*, \tau) = \lambda_i^* \times [1 - \frac{1}{\tau}]^{1/\tau}$, for the WEI5 model. Considering the partitions $\mathbf{x}_i^\top = (1, \mathbf{x}_i^{\Delta\top})$, $\boldsymbol{\beta}^\top = (\beta_0, \boldsymbol{\beta}^{\Delta\top})$ and $\boldsymbol{\delta}^\top = (\delta_0, \boldsymbol{\delta}^{\Delta\top})$, i.e., $\mathbf{x}_i^{\Delta\top}$, $\boldsymbol{\beta}^{\Delta\top}$ and $\boldsymbol{\delta}^{\Delta\top}$ are related to the non-intercept terms of \mathbf{x}_i^\top , $\boldsymbol{\beta}^\top$ and $\boldsymbol{\delta}^\top$, respectively. With this, the equation $\lambda_i = k(\lambda_i^*, \tau)$ assumes the following forms for each case:

- WEI2 model:

$$\exp(-\beta_0 - \mathbf{x}_i^{\Delta\top} \boldsymbol{\beta}^\Delta) = \exp\left(\frac{\delta_0}{\tau} + \frac{\mathbf{x}_i^{\Delta\top} \boldsymbol{\delta}^\Delta}{\tau}\right).$$

In this case, we conclude that $\delta_0 = -\tau\beta_0$ and $\boldsymbol{\beta}^\Delta = -\boldsymbol{\delta}^\Delta/\tau$.

- WEI3 model:

$$\exp(-\beta_0 - \mathbf{x}_i^{\Delta\top} \boldsymbol{\beta}^\Delta) = \exp\left(\delta_0 - \log \tau - \tau \log \Gamma(1 + 1/\tau) + \mathbf{x}_i^{\Delta\top} \boldsymbol{\delta}^\Delta\right).$$

Here, it can be concluded that $\delta_0 = -\beta_0 + \log \tau + \tau \log \Gamma(1 + 1/\tau)$ and $\boldsymbol{\beta}^\Delta = \boldsymbol{\delta}^\Delta$.

- WEI4 model:

$$\exp(-\beta_0 - \mathbf{x}_i^{\Delta\top} \boldsymbol{\beta}^\Delta) = \exp\left(\delta_0 + \log(-\log(1 - q)) + \mathbf{x}_i^{\Delta\top} \boldsymbol{\delta}^\Delta\right),$$

where we conclude that $\delta_0 = -\beta_0 - \log(-\log(1 - q))$ and $\boldsymbol{\beta}^\Delta = \boldsymbol{\delta}^\Delta$.

- WEI5 model:

$$\exp(-\beta_0 - \mathbf{x}_i^{\Delta\top} \boldsymbol{\beta}^\Delta) = \exp\left(\delta_0 + \frac{1}{\tau} \log\left(1 - \frac{1}{\tau}\right) + \mathbf{x}_i^{\Delta\top} \boldsymbol{\delta}^\Delta\right).$$

In this case, we conclude that $\delta_0 = -\beta_0 - \frac{1}{\tau} \log\left(1 - \frac{1}{\tau}\right)$ and $\beta^\Delta = \delta^\Delta$.

Note that in all cases, (δ^\top, τ) can be obtained uniquely from (β^\top, ν) . \square

Remark 2. Theorem 1 implies that when only λ is modeled in the WEI, WEI2, WEI3, WEI4, and WEI5 models, all the distributions produce the same PDF. Consequently, model selection criteria, such as the Akaike Information criterion (AIC) and Bayesian information criterion (BIC), yield identical values.

Remark 3. When an additional set of covariates, denoted as \mathbf{z}_i (not necessarily identical to \mathbf{x}_i), is considered to model the shape parameter as $\nu_i = h(\mathbf{z}_i^\top \boldsymbol{\zeta})$, where $h : \mathbb{R} \rightarrow \mathbb{R}^+$, the five parameterizations of the Weibull model define different PDFs. This distinction arises because the relationship $\lambda_i = k(\lambda_i^*, \tau_i)$ incorporates only the covariates \mathbf{x}_i on the left side of the equation, while both sets of covariates are involved on the right side.

4. Inference

In this section, we discuss the estimation procedure in a unified way for the WEI, WEI2, WEI3, WEI4, and WEI5 models from a classical approach. The asymptotic distribution of the estimators is also presented.

Parameter Estimation

Let Y_1, \dots, Y_n be n independent random variables, where each $Y_i, i = 1, \dots, n$, follows a WEI, WEI2, WEI3, WEI4, or WEI5 model. The log-likelihood function for $\boldsymbol{\theta}^\top = (\beta^\top, \delta^\top)$ has the form

$$\ell(\boldsymbol{\theta}) = \sum_{i=1}^n \ell_i(\lambda_i, \nu_i; y_i),$$

where

$$\ell_i(\lambda_i, \nu_i; y_i) = \log \nu_i - \log \lambda_i + (\nu_i - 1)[\log t_i - \log \lambda_i] - \left[\frac{t_i}{\lambda_i}\right]^{\nu_i},$$

with $\lambda_i = k(\lambda_i^*, \nu_i)$, $\lambda_i^* = g(\mathbf{x}_i^\top \boldsymbol{\beta})$ and $\nu_i = h(\mathbf{z}_i^\top \boldsymbol{\delta})$. The function $k(\cdot, \cdot)$ is defined in Table 2 for the different parameterizations of the WEI model.

Table 2. Function $k(\lambda, \nu)$ for the different parameterizations of the WEI model.

Model	WEI	WEI2	WEI3	WEI4	WEI5
$k(\lambda, \nu)$	λ	$\lambda^{-1/\nu}$	$\lambda[\nu\{\Gamma(1 + 1/\nu)\}^\nu]^{-1}$	$-\lambda \log(1 - q)$	$\lambda \left[1 - \frac{1}{\nu}\right]^{1/\nu}$

The maximum likelihood (ML) estimates of $\boldsymbol{\beta}$ and $\boldsymbol{\delta}$ are computed as the solution of the non-linear system $\mathbf{U}(\boldsymbol{\theta}) = \mathbf{0}_{q+r}$, where $\mathbf{0}_{q+r}$ denotes a $(q + r) \times 1$ vector of zeros and $\mathbf{U}(\boldsymbol{\theta})$ is given by

$$\mathbf{U}(\boldsymbol{\theta}) = \begin{bmatrix} \partial \ell(\boldsymbol{\theta}) / \partial \boldsymbol{\beta} \\ \partial \ell(\boldsymbol{\theta}) / \partial \boldsymbol{\delta} \end{bmatrix} = \begin{bmatrix} \mathbf{X}^\top \mathbf{A}_\lambda \mathbf{C}_\lambda \mathbf{D}_\lambda \\ \mathbf{Z}^\top \mathbf{A}_\nu \mathbf{D}_\nu \end{bmatrix},$$

where $\mathbf{A}_u = \text{diag}(a_u^{(1)}, \dots, a_u^{(n)})$, $\mathbf{C}_\lambda = \text{diag}(c_\lambda^{(1)}, \dots, c_\lambda^{(n)})$, $\mathbf{D}_u = \text{diag}(d_u^{(1)}, \dots, d_u^{(n)})$, for $u \in \{\lambda, \nu\}$, $a_\lambda^{(i)} = g'(\mathbf{x}_i^\top \boldsymbol{\beta})$, $a_\nu^{(i)} = h'(\mathbf{z}_i^\top \boldsymbol{\delta})$ and

$$c_{\lambda}^{(i)} = \begin{cases} 1 & , \text{ for the WEI model} \\ -\frac{1}{\nu} \lambda^{-1/\nu-1} & , \text{ for the WEI2 model} \\ [v\{\Gamma(1 + 1/\nu)\}]^{-1} & , \text{ for the WEI3 model} \\ -\log(1 - q) & , \text{ for the WEI4 model} \\ [1 - 1\nu]^{1/\nu} & , \text{ for the WEI5 model} \end{cases}$$

$$d_{\lambda}^{(i)} = -\frac{v_i}{\lambda_i} + \frac{v_i}{\lambda_i} \left[\frac{t_i}{\lambda_i} \right]^{v_i},$$

$$d_{\nu}^{(i)} = \frac{1}{v_i} + \log t_i - \log \lambda_i - \left[\frac{t_i}{\lambda_i} \right]^{v_i} [\log t_i - \log \lambda_i].$$

The Fisher information matrix for models can be written as

$$\mathbf{I}(\boldsymbol{\theta}) = \mathbb{E} \begin{bmatrix} -\partial^2 \ell(\boldsymbol{\theta}) / \partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^{\top} & -\partial^2 \ell(\boldsymbol{\theta}) / \partial \boldsymbol{\beta} \partial \boldsymbol{\delta}^{\top} \\ \cdot & -\partial^2 \ell(\boldsymbol{\theta}) / \partial \boldsymbol{\delta} \partial \boldsymbol{\delta}^{\top} \end{bmatrix} = \tilde{\mathbf{X}}^{\top} \mathbf{W}(\boldsymbol{\theta}) \tilde{\mathbf{X}},$$

where

$$\tilde{\mathbf{X}} = \begin{pmatrix} \mathbf{X} & \mathbf{0}_{p \times 1} \\ \mathbf{0}_{q \times 1} & \mathbf{Z} \end{pmatrix} \quad \text{and}$$

$$\mathbf{W}(\boldsymbol{\theta}) = \begin{bmatrix} \mathbf{A}_{\lambda} \mathbf{C}_{\lambda} \mathbf{V}_{\lambda\lambda} \mathbf{C}_{\lambda} \mathbf{A}_{\lambda} & \mathbf{A}_{\lambda} \mathbf{V}_{\lambda\nu} \mathbf{C}_{\lambda} \mathbf{A}_{\nu} \\ \mathbf{A}_{\nu} \mathbf{V}_{\lambda\nu} \mathbf{C}_{\lambda} \mathbf{A}_{\lambda} & \mathbf{A}_{\nu} \mathbf{V}_{\nu\nu} \mathbf{A}_{\nu} \end{bmatrix},$$

with $\mathbf{V}_{uv} = \text{diag}(V_{uv}^{(1)}, \dots, V_{uv}^{(n)})$, for $u, v \in \{\lambda, \nu\}$, where

$$V_{\lambda\lambda}^{(i)} = -\mathbb{E} \left(\frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \lambda_i^2} \right) = -\mathbb{E} \left(\frac{v_i}{\lambda_i^2} - \frac{v_i(v_i + 1)}{\lambda_i^2} \left[\frac{t_i}{\lambda_i} \right]^{v_i} \right) = \frac{v_i^2}{\lambda_i^2},$$

$$V_{\nu\nu}^{(i)} = -\mathbb{E} \left(\frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial v_i^2} \right) = -\mathbb{E} \left(-\frac{1}{v_i^2} - \left[\frac{t_i}{\lambda_i} \right]^{v_i} [\log t_i - \log \lambda_i] \right)$$

$$= \frac{1}{v_i^2} + \frac{1}{v_i} [\psi'(1) + \psi^2(1) + 2v_i \log \lambda_i \psi(1) + v_i^2 \log^2 \lambda_i - \psi(2) - v_i \log \lambda_i + v_i \log^2 \lambda_i],$$

$$V_{\lambda\nu}^{(i)} = -\mathbb{E} \left(\frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \lambda_i \partial v_i} \right) = -\mathbb{E} \left(-\frac{1}{\lambda_i^2} + \frac{1}{\lambda_i} \left[\frac{t_i}{\lambda_i} \right]^{v_i} + \frac{v_i}{\lambda_i} \left[\frac{t_i}{\lambda_i} \right]^{v_i} [\log t_i - \log \lambda_i] \right) = -\frac{\psi(2)}{\lambda_i}.$$

An alternative way to obtain the ML estimates of $\boldsymbol{\theta}$ is by using the Fisher scoring iterative procedure, providing the following estimation algorithm:

$$\hat{\boldsymbol{\theta}}^{(k+1)} = \hat{\boldsymbol{\theta}}^{(k)} + [\mathbf{I}(\hat{\boldsymbol{\theta}}^{(k)})]^{-1} \times \mathbf{U}(\hat{\boldsymbol{\theta}}^{(k)})$$

$$= \hat{\boldsymbol{\theta}}^{(k)} + [\tilde{\mathbf{X}}^{\top} \mathbf{W}(\hat{\boldsymbol{\theta}}^{(k)}) \tilde{\mathbf{X}}]^{-1} \tilde{\mathbf{X}}^{\top} \mathbf{W}_1(\hat{\boldsymbol{\theta}}^{(k)}), \quad k = 0, 1, 2, \dots,$$

where $\mathbf{W}_1(\boldsymbol{\theta}) = (\mathbf{A}_{\lambda} \mathbf{C}_{\lambda} \mathbf{D}_{\lambda}, \mathbf{A}_{\nu} \mathbf{D}_{\nu})^{\top}$. We recommend initializing the algorithm by using, as an initial guess for $\boldsymbol{\beta}$, the ordinary least squares estimates obtained from the linear regression of the transformed responses: $g(y_1), \dots, g(y_n)$ on \mathbf{X} , i.e., $(\mathbf{X}^{\top} \mathbf{X})^{-1} \mathbf{X}^{\top} \mathbf{Z}$, where $\mathbf{Z} = (g(y_1), \dots, g(y_n))^{\top}$.

In order to approach the interval estimation and hypothesis testing on the model parameters $\boldsymbol{\beta}$ and $\boldsymbol{\delta}$, normal approximation for the ML estimators can be applied. Note that under certain conditions for the parameters, the asymptotic distribution of $\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})$ is multivariate normal $N_{q+r}(\mathbf{0}_{q+r}, \mathbf{I}^{-1}(\boldsymbol{\theta}))$.

Let the θ_j be the j th component of θ . The asymptotic $100(1 - \gamma)\%$ confidence interval for θ_j is given by

$$\hat{\theta}_j \pm z_{1-\gamma/2} \text{se}(\hat{\theta}_j), \quad j = 1, \dots, q + r,$$

where $\text{se}(\hat{\theta}_j)$ is the asymptotic standard error of $\hat{\theta}_j$, which is the square root of the j th diagonal element of the inverse of $\mathbf{I}(\hat{\theta})$.

5. Monte Carlo Simulation Studies

In this section, we assess the performances of the ML estimators through MC simulations, compare their empirical biases and mean squared error (MSE), and evaluate the residuals for WEI models. The WEI4 and WEI5 models are implemented in R software (version 4.2, accessed on 5 December 2023). www.r-project.org and a routine has been developed for the aforementioned simulations. The R-code can be obtained from https://github.com/lsanchez2020/Weibull_reparameterized.git.

5.1. MC Simulation Study I: Behavior of the ML Estimators

For WEI4 (with $q = 0.5$) and WEI5 models, the following regression structure is considered:

$$\lambda_i = \exp(\mathbf{x}_i^\top \boldsymbol{\beta}), \quad \text{and} \quad v_i = \exp(\mathbf{z}_i^\top \boldsymbol{\zeta}), \quad i = 1, \dots, n.$$

To assess the performance of the ML estimators for the WEI4 and WEI5 models, both the bias and the MSE are reported. The bias and MSE are calculated as follows:

$$\widehat{\text{Bias}}(\hat{\varphi}) = \frac{1}{B} \sum_{j=1}^B (\hat{\varphi}_j - \varphi), \quad \widehat{\text{MSE}}(\hat{\varphi}) = \frac{1}{B} \sum_{j=1}^B (\hat{\varphi}_j - \varphi)^2,$$

where $\hat{\varphi}_j$ is the estimate in the j th replicate, φ is the true value of the parameter, and B is the number of MC replicates; 5000 MC replicates are conducted, where the explanatory variables are simulated from a Uniform (0, 1) distribution. Four scenarios for the true values of the parameters are considered: (i) true values of $\beta_1, \beta_2, \zeta_1$, and ζ_2 are 0.2, 1.0, 0.2, and 1.0; (ii) true values of $\beta_1, \beta_2, \zeta_1$, and ζ_2 are 0.2, 0.5, 0.2, and 1.0; (iii) true values of $\beta_1, \beta_2, \zeta_1$, and ζ_2 are 0.5, 1.0, 0.2, and 1.0; and (iv) true values of $\beta_1, \beta_2, \zeta_1$, and ζ_2 are 0.5, 1.5, 0.2, and 1.0.

Tables 3 and 4 report the ML estimation results for the bias and the MSE of the WEI4 and WEI5 model estimators, respectively. A general trend observed in these tables is that—as the sample size increases—both bias and MSE decrease, as expected in a Monte Carlo study. Therefore, the results confirm the consistence and un-biasing of the ML estimators.

Table 3. Bias and MSE for the WEI4 model parameters, with the indicated sample size and true values of the parameters.

$\beta_1 = 0.2; \beta_2 = 1.0; \zeta_1 = 0.2; \zeta_2 = 1.0$								
Bias					MSE			
n	β_1	β_2	ζ_1	ζ_2	β_1	β_2	ζ_1	ζ_2
30	−0.0028	0.0013	0.0761	0.0232	0.0808	0.2592	0.2330	0.6466
50	−0.0039	0.0035	0.0371	0.0296	0.0502	0.1470	0.0972	0.3091
100	−0.0005	0.0007	0.0152	0.0179	0.0206	0.0661	0.0411	0.1357
500	−0.0001	0.0000	0.0048	−0.0005	0.0043	0.0134	0.0080	0.0253

Table 3. Cont.

$\beta_1 = 0.2; \beta_2 = 0.5; \zeta_1 = 0.2; \zeta_2 = 1.0$								
Bias					MSE			
n	β_1	β_2	ζ_1	ζ_2	β_1	β_2	ζ_1	ζ_2
30	−0.0153	0.0185	0.0533	0.0503	0.0843	0.2499	0.1846	0.5915
50	0.0021	−0.0036	0.0368	0.0318	0.0430	0.1333	0.0937	0.3019
100	0.0013	−0.0010	0.0144	0.0223	0.0206	0.0619	0.0470	0.1478
500	−0.0004	0.0001	0.0037	0.0029	0.0041	0.0134	0.0075	0.0238
$\beta_1 = 0.5; \beta_2 = 1.0; \zeta_1 = 0.2; \zeta_2 = 1.0$								
Bias					MSE			
n	β_1	β_2	ζ_1	ζ_2	β_1	β_2	ζ_1	ζ_2
30	0.0024	−0.0136	0.0596	0.0539	0.0867	0.2542	0.1826	0.5722
50	−0.0047	0.0017	0.0326	0.0285	0.0515	0.1707	0.0911	0.3315
100	−0.0004	−0.0012	0.0184	0.0132	0.0238	0.0713	0.0403	0.1326
500	−0.0016	0.0023	0.0020	0.0038	0.0045	0.0137	0.0078	0.0248
$\beta_1 = 0.5; \beta_2 = 1.5; \zeta_1 = 0.2; \zeta_2 = 1.0$								
Bias					MSE			
n	β_1	β_2	ζ_1	ζ_2	β_1	β_2	ζ_1	ζ_2
30	0.0007	−0.0067	0.0533	0.0574	0.0802	0.2363	0.1851	0.6371
50	−0.0052	0.0021	0.0315	0.0368	0.0487	0.1529	0.0943	0.3092
100	−0.0012	0.0016	0.0194	0.0071	0.0251	0.0714	0.0422	0.1316
500	0.0006	−0.0020	0.0040	0.0022	0.0045	0.0134	0.0076	0.0243

Table 4. Bias and MSE for the WEI5 model parameters, with the indicated sample size and true values of the parameters.

$\beta_1 = 0.2; \beta_2 = 1.0; \zeta_1 = 0.2; \zeta_2 = 1.0$								
Bias					MSE			
n	β_1	β_2	ζ_1	ζ_2	β_1	β_2	ζ_1	ζ_2
30	−0.0019	0.0099	0.0192	0.1226	0.1699	0.4900	0.0889	0.3378
50	0.0082	−0.0045	0.0149	0.0620	0.0877	0.2437	0.0454	0.1648
100	0.0075	−0.0051	0.0050	0.0333	0.0659	0.1702	0.0158	0.0642
500	0.0003	0.0001	0.0031	0.0022	0.0067	0.0177	0.0025	0.0110
$\beta_1 = 0.2; \beta_2 = 0.5; \zeta_1 = 0.2; \zeta_2 = 1.0$								
Bias					MSE			
n	β_1	β_2	ζ_1	ζ_2	β_1	β_2	ζ_1	ζ_2
30	0.0043	−0.0059	0.0211	0.1135	0.2746	0.5124	0.0865	0.3565
50	0.0021	0.0046	0.0164	0.0617	0.1146	0.2692	0.0366	0.1447
100	0.0012	0.0027	0.0035	0.0396	0.0454	0.1138	0.0178	0.0673
500	0.0015	−0.0020	0.0018	0.0058	0.0067	0.0181	0.0027	0.0108
$\beta_1 = 0.5; \beta_2 = 1.0; \zeta_1 = 0.2; \zeta_2 = 1.0$								
Bias					MSE			
n	β_1	β_2	ζ_1	ζ_2	β_1	β_2	ζ_1	ζ_2
30	0.0074	−0.0023	0.0188	0.1126	0.1963	0.5392	0.0832	0.3169
50	−0.0019	0.0055	0.0123	0.0638	0.1380	0.3698	0.0434	0.1694
100	0.0006	−0.0019	0.0035	0.0338	0.0360	0.1027	0.0158	0.0626
500	0.0015	−0.0020	0.0009	0.0076	0.0066	0.0175	0.0026	0.0107

Table 4. Cont.

$\beta_1 = 0.5; \beta_2 = 1.5; \zeta_1 = 0.2; \zeta_2 = 1.0$								
Bias					MSE			
n	β_1	β_2	ζ_1	ζ_2	β_1	β_2	ζ_1	ζ_2
30	−0.0136	0.0217	0.0205	0.1142	0.4725	1.5218	0.0842	0.3431
50	0.0013	0.0040	0.0104	0.0720	0.0886	0.2640	0.0414	0.1756
100	−0.0034	0.0078	0.0056	0.0324	0.0485	0.1301	0.0226	0.0738
500	0.0004	−0.0003	0.0010	0.0063	0.0062	0.0169	0.0027	0.0108

5.2. MC Simulation Study II: Residuals Analysis

We delve into the analysis of the Cox–Snell (CS) and randomized quantile (RQ) residuals for the WEI4 and WEI5 models. The computation of the CS residual is as follows:

$$\hat{r}_i^{CS} = -\log(\hat{S}(y_i; x_i)), \quad i = 1, \dots, n,$$

where $\hat{S}(y_i|x_i)$ is the estimated survival function obtained from the fitted model with values x_i for the explanatory variables. The survival functions of the models under consideration are presented in Section 2. The RQ residual is defined as

$$\hat{r}_i^{RQ} = \Phi^{-1}(\hat{S}(y_i|x_i)), \quad i = 1, \dots, n,$$

where Φ^{-1} is the quantile function of the standard normal distribution.

For this investigation, we maintain the same simulation scenario as in the MC simulation study I, as outlined above. Tables 5 and 6 display the empirical mean, standard deviation (SD), and coefficient of skewness (CSk) of the CS and RQ residuals. Theoretical expectations for these statistics are 1, 1, and 2, respectively, for the CS residuals, and 0, 1, and 0, respectively, for the RQ residuals, due to their respective theoretical distributions being exponential and standard normal. The results in Tables 5 and 6 demonstrate that, in general, both types of residuals provide good approximations for the WEI4 and WEI5 models.

Table 5. Mean, SD, and CSk of the CS and randomized quantile residuals for the WEI4 model, with the indicated sample size and true values of the parameters.

$\beta_1 = 0.2; \beta_2 = 1.0; \zeta_1 = 0.2; \zeta_2 = 1.0$						
Randomized Quantile				Cox–Snell		
n	Mean	SD	CSk	Mean	SD	CSk
30	−0.0176	1.0197	0.0225	0.9879	0.9686	1.2599
50	−0.0064	1.0067	0.0761	0.9936	1.0119	1.8385
100	0.0057	0.9918	0.1559	1.0024	1.0380	1.9373
500	0.0085	0.9903	0.1282	1.0057	1.0256	1.9420
$\beta_1 = 0.2; \beta_2 = 0.5; \zeta_1 = 0.2; \zeta_2 = 1.0$						
randomized quantile				Cox–Snell		
n	mean	SD	CSk	mean	SD	CSk
30	−0.0124	0.9931	0.3132	0.9862	1.0388	1.4458
50	−0.0045	1.0443	−0.3680	1.0011	0.9207	1.3308
100	−0.0268	1.0082	−0.0545	0.9732	1.0482	2.7658
500	−0.0038	0.9844	0.2264	0.9942	1.0366	2.0337

Table 5. Cont.

$\beta_1 = 0.5; \beta_2 = 1.0; \zeta_1 = 0.2; \zeta_2 = 1.0$						
randomized quantile				Cox–Snell		
n	mean	SD	CSk	mean	SD	CSk
30	−0.0190	1.0429	−0.2383	0.9903	0.9126	1.0349
50	−0.0207	1.0441	−0.4426	0.9851	0.9050	1.1128
100	−0.0300	1.0318	−0.2247	0.9815	0.9164	1.8438
500	−0.0064	1.0036	−0.0312	0.9937	1.0155	2.4167
$\beta_1 = 0.5; \beta_2 = 1.5; \zeta_1 = 0.2; \zeta_2 = 1.0$						
randomized quantile				Cox–Snell		
n	mean	SD	CSk	mean	SD	CSk
30	0.0280	0.9916	0.3017	1.0226	1.0364	1.2727
50	0.0098	1.0222	−0.1606	1.0113	0.9557	1.2558
100	−0.0207	1.0208	−0.1657	0.9843	0.9607	1.6202
500	−0.0116	0.9978	0.0671	0.9904	0.9961	1.9101

Table 6. Mean, SD, and CSk of the Cox–Snell and randomized quantile residuals for the WEI5 model, with the indicated sample size and true values of the parameters.

$\beta_1 = 0.2; \beta_2 = 1.0; \zeta_1 = 0.2; \zeta_2 = 1.0$						
Randomized Quantile				Cox–Snell		
n	Mean	SD	CSk	Mean	SD	CSk
30	0.0319	0.9772	0.0273	1.0073	0.9587	1.3523
50	−0.0288	1.0180	0.1452	0.9868	0.9880	1.1914
100	0.0152	1.0165	−0.3226	1.0084	0.9403	1.4739
500	−0.0029	1.0066	−0.0848	0.9978	0.9802	1.7448
$\beta_1 = 0.2; \beta_2 = 0.5; \zeta_1 = 0.2; \zeta_2 = 1.0$						
randomized quantile				Cox–Snell		
n	mean	SD	CSk	mean	SD	CSk
30	0.0726	0.9760	−0.0228	1.0433	0.9477	1.2336
50	0.0131	1.0012	−0.0942	1.0043	0.9477	1.5201
100	−0.0237	0.9867	0.3180	0.9791	1.0466	2.1566
500	0.0099	0.9854	0.1456	1.0048	1.0183	2.0615
$\beta_1 = 0.5; \beta_2 = 1.0; \zeta_1 = 0.2; \zeta_2 = 1.0$						
randomized quantile				Cox–Snell		
n	mean	SD	CSk	mean	SD	CSk
30	−0.0715	1.0858	−0.1631	0.9710	0.9773	1.7365
50	−0.0193	1.0235	−0.0200	0.9914	0.9712	1.6429
100	−0.0115	1.0165	−0.0516	0.9960	0.9590	1.4681
500	−0.0127	1.0102	−0.0333	0.9930	0.9906	1.8522
$\beta_1 = 0.5; \beta_2 = 1.5; \zeta_1 = 0.2; \zeta_2 = 1.0$						
randomized quantile				Cox–Snell		
n	mean	SD	CSk	mean	SD	CSk
30	0.0197	1.0379	−0.4223	1.0133	0.9251	1.3285
50	0.0156	1.0119	0.0220	1.0156	1.0344	2.2866
100	0.0032	1.0122	−0.1720	1.0023	0.9195	1.1461
500	0.0031	1.0092	−0.1179	1.0044	0.9577	1.5290

6. Real Data Illustrations

In this section, the analyzed parametrizations to real data are applied to demonstrate their effectiveness and applicability. In Illustrations 1 and 2, models WEI4 and WEI5, as well as WEI1, WEI2, and WEI3, respectively, showcase their applicability in real-life scenarios.

6.1. Illustration 1: Concrete Fatigue Life Dataset

The dataset under analysis represents the fatigue life of concrete specimens, measured in cycles (multiplied by 10^{-3}). The dataset is defined by the applied stress-causing failure, with ratios of 0.95, 0.90, and 0.825. A lower ratio is expected to correspond to a longer number of cycles until failure. Each ratio is accompanied by 15 observations, as outlined in ([46], p. 88). The concrete fatigue life is designated as the response variable, and the ratio is identified as the covariate.

Initially, we provide an exploratory analysis of the response variable. Table 7 presents some descriptive statistics, including SD, CSk, the coefficient of variation (CV), and coefficient of kurtosis (CK). Additionally, Figure 1 displays the corresponding histogram and boxplot. From Table 7 and Figure 1, it is evident that the dataset exhibits considerable variability (CV = 131.401%), positive skewness (CSk = 1.414), heavy tails (CK = 0.771), and the presence of outliers (case nos. 40, 41, 42, 43, 44, and 45).

Table 7. Descriptive statistics of the concrete fatigue life dataset.

<i>n</i>	min	Max	Mean	Median	SD	CSk	CV	CK
45	37	5598	1216.58	342	1598.60	1.41	131.40	0.77

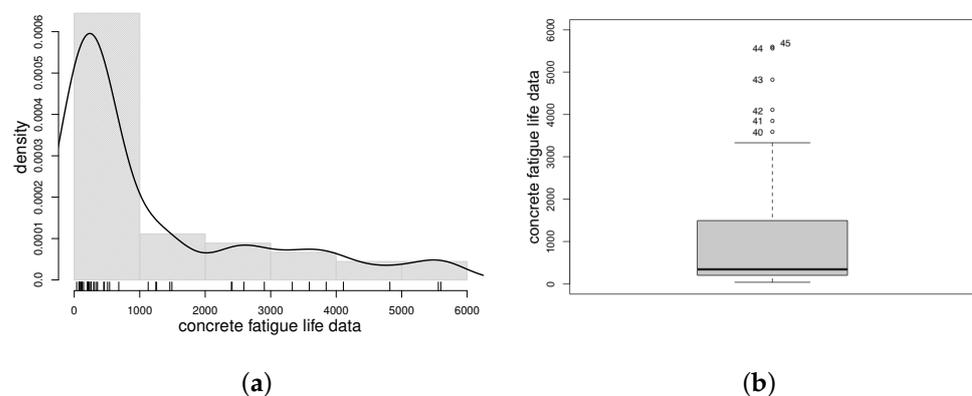


Figure 1. Histogram with the density kernel estimation (a) and boxplot (b) for the concrete fatigue life dataset.

In this study, a regression structure is incorporated into the WEI4 (with $q = 0.5$) and WEI5 models for both λ and ν , as follows:

$$\lambda_i = \exp(\mathbf{x}_i^\top \boldsymbol{\beta}), \quad \nu_i = \exp(\mathbf{x}_i^\top \boldsymbol{\zeta}), \quad i = 1, \dots, 45,$$

where $\boldsymbol{\beta} = (\beta_0, \beta_1)^\top$, and $\boldsymbol{\zeta} = (\zeta_0, \zeta_1)^\top$ represent the regression coefficients, and $\mathbf{x}_i^\top = (1, x_{1i})$ denotes the respective design matrix, with x_{1i} being the i -th observation of the covariate ratio.

Table 8 presents the fitting of these models, including parameter estimation using the ML method, the respective standard errors, the associated Z-test p -value, the AIC, and the p -value from the Kolmogorov–Smirnov (KS), Anderson–Darling (AD), and Cramer von Mises (CVM) goodness-of-fit tests for the CS and RQ residuals. It is important to note that the theoretical distributions for CS and RQ residuals are exponential and standard normal, respectively. In this table, it can be observed that for modeling λ , both regression

coefficients (intercept and slope) are significant. However, this situation is not observed in the modeling of ν , where none of the regression coefficients are significant.

Additionally, Figures 2 and 3 provide quantile–quantile (QQ) goodness-of-fit plots with simulated envelopes for the CS and RQ residuals. Consistent with the goodness-of-fit test results, both residuals are well-fitted to the exponential and standard normal distributions, respectively.

Table 8. ML estimates, SEs, p -values, AIC, and KS p -value of the indicated model.

		WEI4			WEI5		
	Parameter	Estimate	SE	p -Value	Estimate	SE	p -Value
λ	β_1	29.2805	1.4737	<0.0001	29.5963	1.8853	<0.0001
	β_2	−25.8166	1.6582	<0.0001	−26.3132	2.1276	<0.0001
ν	ζ_1	1.4702	2.2139	0.5066	1.5362	1.8117	0.3965
	ζ_2	−0.7999	2.4753	0.7466	−0.8740	2.0218	0.6655
AIC		635.03			635.02		
KS p -value		0.8729 (CS)	0.4339 (RQ)		0.8753 (CS)	0.4102 (RQ)	
AD p -value		—	0.1082 (RQ)		—	0.1117 (RQ)	
CVM p -value		—	0.1203 (RQ)		—	0.1234 (RQ)	

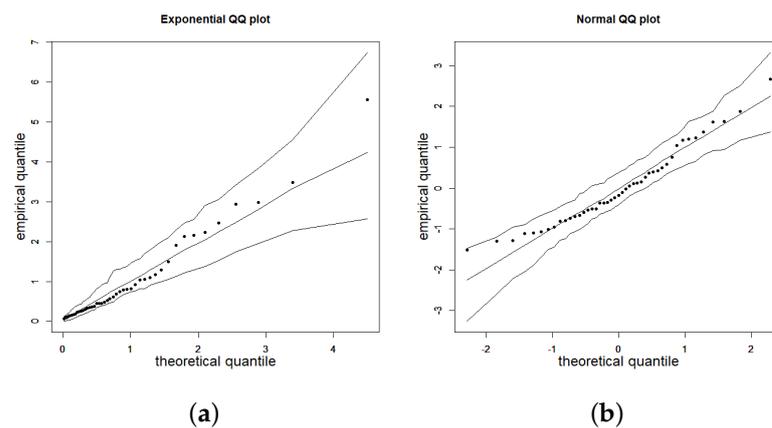


Figure 2. QQ plot with the envelope of CS (a) and RQ (b) residuals for the WEI4 model.

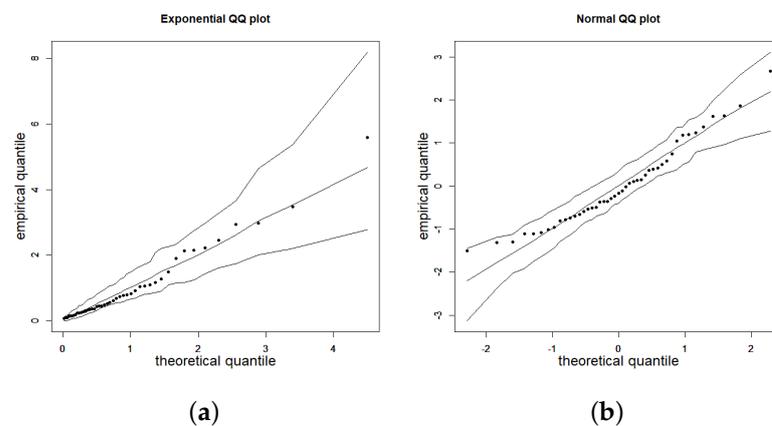


Figure 3. QQ plot with the envelope of CS (a) and RQ (b) residuals for the WEI5 model.

6.2. Illustration 2: Ovarian Dataset

The ovarian cancer data (ovarian) are also considered, which are included in the survival package, R version 3.5-7 [47]. This dataset contains censored observations and

will be used to illustrate the reparameterized Weibull model in the context of censored data. It comprises information on 26 patients with stage III and IV ovarian carcinoma who were treated with radical surgery and postoperative chemotherapy. The covariates used are the minimal residual disease (*residual*), which indicates a tumor mass of less than 2 cm in diameter, and the patient’s age in years. Figure 4 displays the Kaplan–Meier (K-M) estimator for both *residual* and age. Descriptively, it is evident that both covariates have an impact on patient survival. In addition, and similar to the last application, in Figure 5, the QQ plot and simulated envelope for the RQ residuals are provided, suggesting that this kind of residual is reasonably fitted to the standard normal distribution.

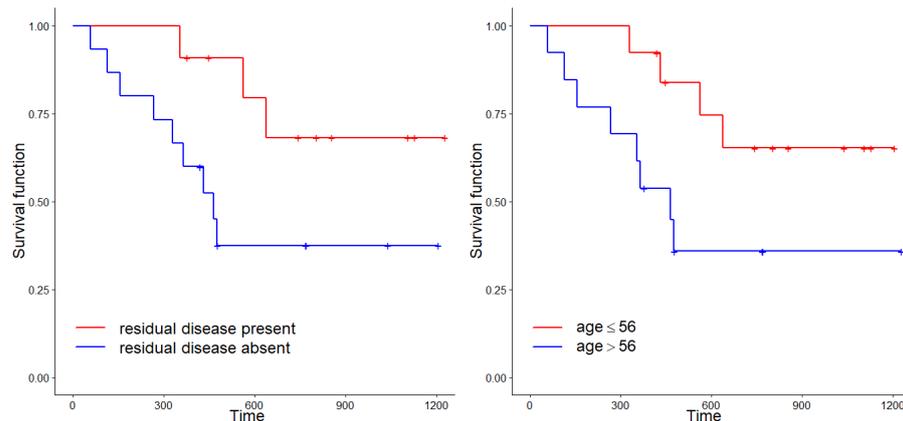


Figure 4. K-M estimator for ovarian dataset separated by the *residual* (left panel) and age (right panel).

Table 9 shows the AIC for the WEI, WEI2, and WEI3 models in the ovarian dataset with different combinations of covariates in λ and ν . As discussed previously, the AIC is the same for the three models if covariates are not included in ν and differ if ν is modeled. The results also suggest that the best combination is to include the covariates *residual* plus age in λ , but not in ν . Table 10 shows the estimate, the corresponding SE for the WEI2 model (it could have been any of them), and the p -value from the KS, AD, and CVM goodness-of-fit tests for RQ residuals.

Table 9. The AIC for the fitted model ovarian dataset.

λ	Covariates		WEI	Model	
		ν		WEI2	WEI3
Only intercept		Only intercept		199.91	
Residual plus age		Only intercept		186.04	
Residual plus age		Residual plus age	189.56	190.04	189.53

Finally, Figure 6 shows the mean, median, and mode of the WEI, WEI2, and WEI3 models. As proved in Theorem 1, such measures are identical for the three models because no covariates are included in the dispersion parameter.

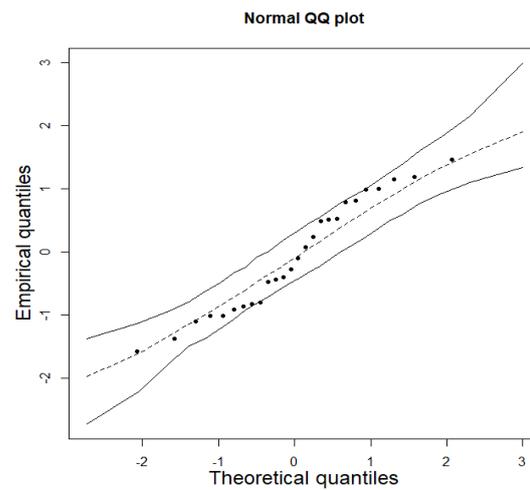


Figure 5. QQ plot with the envelope of RQ residuals for the ovarian dataset. Note that the plot is identical for WEI, WEI2, and WEI3 models.

Table 10. ML estimate, standard error (SE) for WEI2, considering the residual plus age covariates for the ovarian dataset.

		WEI2	
	Parameters	Estimate	SE
λ	Intercept	−20.5374	1.1339
	residual	0.9720	0.4152
	age	0.1401	0.0207
ν	Intercept	0.5559	0.0155
AIC		186.04	
KS <i>p</i> -value		0.6703	
AD <i>p</i> -value		0.1916	
CVM <i>p</i> -value		0.1895	

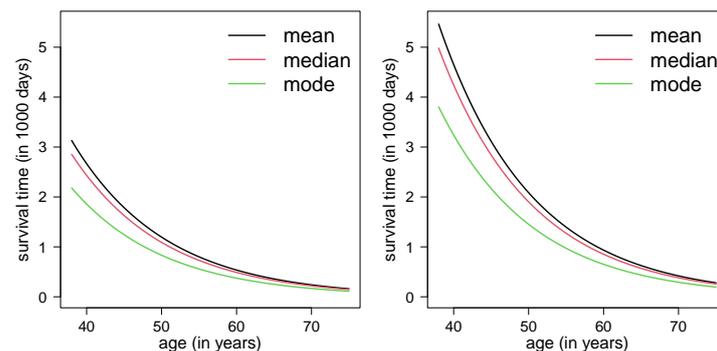


Figure 6. Different measures of central tendency for the survival time in the ovarian problem in terms of age: residual = 1 (left panel) and residual = 2 (right panel). The plots are identical for WEI, WEI2, and WEI3 models.

7. Conclusions and Future Work

In this paper, a general parameterization for Weibull regression models was introduced, based on central tendency measures and shape parameters. A comprehensive review of Weibull regression models was provided, specifically focusing on different parameterizations of the Weibull distribution that utilize central tendency measures. The interpretation of regression coefficients when incorporating regression structures into these

parameterizations was also explored. Furthermore, closed-form expressions for the expected Fisher information matrix have been derived for this general parameterization of Weibull regression models. Two types of residuals were explored, and maximum likelihood inference was implemented for estimating model parameters. The performance of this inference method was assessed through Monte Carlo simulations.

A significant result, as established in Theorem 1, is that when modeling only λ in the WEI, WEI2, WEI3, WEI4, and WEI5 models, the same probability density function is yielded by all these parameterizations. Therefore, in these cases, it does not matter which model is used, as the conclusions for any measure of interest must be the same for any of the five models.

To illustrate the practical relevance of the approach, two real-world applications using authentic datasets were presented. These applications underscored the adequacy of the introduced Weibull models when data present an asymmetric distribution. Depending on which parameter is of interest in the study, one model will be more convenient to use than the other.

As part of future research, the plan is to develop an R package that facilitates inference in WEI4 and WEI5 regression models.

Author Contributions: Conceptualization, D.I.G., Y.M.G. and M.B.; methodology, D.I.G., Y.M.G. and M.B.; software, D.I.G., Y.M.G., C.M., L.S. and M.B.; validation, formal analysis, D.I.G., Y.M.G., C.M., L.S. and M.B.; investigation, D.I.G., Y.M.G., C.M., L.S. and M.B.; resources; data curation, Y.M.G. and C.M.; writing—original draft preparation, D.I.G., Y.M.G., C.M., L.S. and M.B.; writing—review and editing, D.I.G., Y.M.G., C.M., L.S. and M.B.; visualization, D.I.G., Y.M.G., C.M. and L.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding

Data Availability Statement: Data and computational codes are available upon request from the authors.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Weibull, W. A statistical distribution function of wide applicability. *J. Appl. Mech.* **1951**, *18*, 293–296 [[CrossRef](#)]
- Keshevan, K.; Sargent, G.; Conrad H. Statistical analysis of the Hertzian fracture of pyrex glass using the Weibull distribution function. *J. Mater. Sci.* **1980**, *15*, 839–844. [[CrossRef](#)]
- Queeshi, F.S.; Sheikh, A.K. Probabilistic characterization of adhesive wear in metals. *IEEE Trans. Reliab.* **1997**, *46*, 38–44.
- Durham, S.D.; Padgett, W.J. Cumulative damage model for system failure with application to carbon fibers and composites. *Technometrics* **1997**, *39*, 34–44. [[CrossRef](#)]
- Almeida, J.B. Application of Weibull statistics to the failure of coatings. *J. Mater. Process. Technol.* **1999**, *93*, 257–263. [[CrossRef](#)]
- Fok, S.L.; Mitchell, B.C.; Smart, J.; Marsden, B.J. A numerical study on the application of the Weibull theory to brittle materials. *Eng. Fract. Mech.* **2001**, *68*, 1171–1179. [[CrossRef](#)]
- Newell, J.A.; Kurzeja, T.; Spence, M.; Lynch, M. Analysis of recoil compressive failure in high performance polymers using two-, four-parameter Weibull models. *High Perform. Polym.* **2002**, *14*, 425–434. [[CrossRef](#)]
- Li, Q.S.; Fang, J.Q.; Liu, D.K.; Tang, J. Failure probability prediction of concrete components. *Cem. Concr. Res.* **2003**, *33*, 1631–1636. [[CrossRef](#)]
- Bebbington, M.S.; Lai, C.D. On nonhomogeneous models for volcanic eruptions. *Math. Geol.* **1996**, *28*, 585–599. [[CrossRef](#)]
- Durrans, S.R. Low-flow analysis with a conditional Weibull tail model. *Water Resour. Res.* **1996**, *32*, 1749–1760. [[CrossRef](#)]
- Heo, J.H.; Boes, D.C.; Salas, J.D. Regional flood frequency analysis based on a Weibull model: Part 1. Estimation, asymptotic variances. *J. Hydrol.* **2001**, *242*, 157–170. [[CrossRef](#)]
- Fleming, R.A. The Weibull model and an ecological application: Describing the dynamics of foliage biomass on Scots pine. *Ecol. Model.* **2001**, *138*, 309–319. [[CrossRef](#)]
- Roed, K.; Zhang, T. A note on the Weibull distribution and time aggregation bias. *Appl. Econ. Lett.* **2002**, *9*, 469–472. [[CrossRef](#)]
- Ikki, S.S.; Ahmed, M.H. Performance of multi-hop relaying systems over Weibull fading channels. In *New Technologies, Mobility and Security*; Springer: Amsterdam, The Netherlands, 2007; pp. 31–38.
- Wang, P.; Zhang, J.; Guo, L.; Shang, T.; Cao, T.; Wang, R.; Yang, T. Performance analysis for relay-aided multihop BPPM FSO communication system over exponentiated Weibull fading channels with pointing error impairments. *IEEE Photon. J.* **2015**, *7*, 1–20. [[CrossRef](#)]
- Rinne, H. *The Weibull Distribution: A Handbook*; Chapman and Hall/CRC: New York, NY, USA, 2008.

17. Lai, C.; Xie, M. Weibull Distributions and Their Applications. In *Springer Handbook of Engineering Statistics*; Pham, H., Ed.; Springer: New York, NY, USA, 2006; pp. 63–78.
18. Lawless, J.F. *Statistical Models and Methods for Lifetime Data*; Wiley: New York, NY, USA, 1982.
19. Ghitany, M.E.; Al-Hussaini, E.K.; Al-Jarallah, R.A. Marshall–Olkin extended weibull distribution and its application to censored data. *J. Appl. Stat.* **2005**, *32*, 1025–1034. [[CrossRef](#)]
20. Klakattawi, H.S. Survival analysis of cancer patients using a new extended Weibull distribution. *PLoS ONE* **2022**, *17*, e0264229. [[CrossRef](#)]
21. Joarder, A.; Krishna, H.; Kundu, D. Inferences on Weibull parameters with conventional type-I censoring. *Comput. Stat. Data Anal.* **2011**, *55*, 1–11. [[CrossRef](#)]
22. Jia, X.; Wang, D.; Jiang, P.; Guo, B. Inference on the reliability of Weibull distribution with multiply Type-I censored data. *Reliab. Eng. Syst. Saf.* **2016**, *150*, 171–181. [[CrossRef](#)]
23. Lee, C.; Famoye, F.; Olumolade, O. Beta-Weibull Distribution: Some Properties and Applications to Censored Data. *J. Mod. Appl. Stat. Methods* **2007**, *6*, 173–186. [[CrossRef](#)]
24. Chen, M.H.; Ibrahim, J.G.; Sinha, D. A new Bayesian model for survival data with a surviving fraction. *J. Am. Stat. Assoc.* **1999**, *94*, 909–919. [[CrossRef](#)]
25. Yin, G.; Ibrahim, J.G. Cure rate models: A unified approach. *Can. J. Stat.* **2005**, *33*, 559–570. [[CrossRef](#)]
26. Rodrigues, J.; Cordeiro, G.M.; Cancho, V.; Balakrishnan, N. Relaxed Poisson cure rate models. *Biom. J.* **2016**, *58*, 397–415. [[CrossRef](#)] [[PubMed](#)]
27. Gallardo, D.I.; Gómez, Y.M.; Gómez, H.W.; De Castro, M. On the use of the modified power series family of distributions in a cure rate model context. *Stat. Methods Med. Res.* **2020**, *29*, 1831–1845. [[CrossRef](#)] [[PubMed](#)]
28. Azimi, R.; Esmailian, M.; Gallardo, D.I.; Gómez, H.J. A New Cure Rate Model Based on Flory–Schulz Distribution: Application to the Cancer Data. *Mathematics* **2022**, *10*, 4643. [[CrossRef](#)]
29. Murthy, D.N.P.; Baik, J.; Bulmer, M.; Wilson, R.J. Two-dimensional modelling of failures. In *Springer Handbook of Engineering Statistics*; Pham, H., Ed.; Springer: Heidelberg/Berlin, Germany, 2004.
30. Murthy, D.N.P.; Xie, M.; Jiang, R. *Weibull Models*; Wiley: New York, NY, USA, 2003.
31. 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](#)]
32. Santos-Neto, M.; Bourguignon, M.; Zea, L.M.; Nascimento, A.D.C.; Cordeiro, G.M. The Marshall–Olkin extended Weibull family of distributions. *J. Stat. Distrib. Appl.* **2014**, *1*, 1–24. [[CrossRef](#)]
33. Nascimento, A.D.C.; Bourguignon, M.; Zea, L.M.; Santos-Neto, M.; Silva, R.B.; Cordeiro, G.M. The gamma extended Weibull family of distributions. *J. Stat. Theory Appl.* **2014**, *13*, 1–16. [[CrossRef](#)]
34. Nadarajah, S.; Kotz, S. On some recent modifications of Weibull distribution. *IEEE Trans. Reliab.* **2005**, *54*, 561–562. [[CrossRef](#)]
35. Pham, H.; Lai, C-D. On recent generalizations of the Weibull distribution. *IEEE Trans. Reliab.* **2007**, *56*, 454–458. [[CrossRef](#)]
36. Almalki, S.J.; Nadarajah, S. Modifications of the Weibull distribution: A review. *Reliab. Eng. Syst. Saf.* **2014**, *124*, 32–55. [[CrossRef](#)]
37. Wingo, D.R. The left-truncated Weibull distribution: Theory and computation. *Stat. Pap.* **1989**, *30*, 39–48. [[CrossRef](#)]
38. Zhang, T.; Xie, M. On the upper truncated Weibull distribution and its reliability implications Reliability. *Eng. Syst. Saf.* **2011**, *96*, 194–200. [[CrossRef](#)]
39. McEwen, R.P.; Parresol, B.R. Moment expressions and summary statistics for the complete and truncated Weibull distribution. *Commun. Stat. Theory Methods* **1991**, *20*, 1361–1372. [[CrossRef](#)]
40. Fernandes, F.H.; Lee, H.L.; Bourguignon, M. About Shewhart control charts to monitor the Weibull mean. *Qual. Reliab. Eng. Int.* **2019**, *35*, 2343–2357. [[CrossRef](#)]
41. Sánchez, L.; Leiva, V.; Saulo, H.; Marchant, C.; Sarabia, J. A new quantile regression model and its diagnostic analytics for a weibull distributed response with applications. *Mathematics* **2021**, *9*, 2768. [[CrossRef](#)]
42. Yao, W.; Li, L. A New Regression Model: Modal Linear Regression. *Scand. J. Stat.* **2013**, *41*, 656–671. [[CrossRef](#)]
43. Chen, Y.C. Modal regression using kernel density estimation: A review. *WIREs Comput. Stat.* **2018**, *10*, e1431. [[CrossRef](#)]
44. Bourguignon, M.; Leao, J.; Gallardo, D.I. Parametric modal regression with varying precision. *Biom. J.* **2020**, *62*, 202–220. [[CrossRef](#)]
45. Kleinbaum, D.G.; Klein, M. Survival analysis : A self-learning text. In *Statistics for Biology and Health*, 3rd ed.; Kleinbaum, D.G., Klein, M., Eds.; Springer: New York, NY, USA, 2012.
46. Leiva, V. *The Birnbaum-Saunders Distribution*; Elsevier: Amsterdam, The Netherlands, 2016.
47. Therneau, T.M. A Package for Survival Analysis in R. R Package Version 3.5-7, CRAN. 2023. Available online: <https://CRAN.R-project.org/package=survival> (accessed on 5 December 2023).

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.