symmetry

MDPI

Article

Skewness of Maximum Likelihood Estimators in the Weibull Censored Data

Tiago M. Magalhães [1], Diego I. Gallardo [2] and Héctor W. Gómez [3,*]

1 Department of Statistics, Institute of Exact Sciences, Federal University of Juiz de Fora, Juiz de Fora 36000-000, Brazil; tiago.magalhaes@ice.ufjf.br

2 Departamento de Matemática, Facultad de Ingeniería, Universidad de Atacama, Copiapó 1530000, Chile; diego.gallardo@uda.cl

3 Departamento de Matemáticas, Facultad de Ciencias Básicas, Universidad de Antofagasta, Antofagasta 1240000, Chile

* Correspondence: hector.gomez@uantof.cl

Received: 29 September 2019; Accepted: 21 October 2019; Published: 1 November 2019

Abstract: In this paper, we obtain a matrix formula of order $n^{-1/2}$, where n is the sample size, for the skewness coefficient of the distribution of the maximum likelihood estimators in the Weibull censored data. The present result is a nice approach to verify if the assumption of the normality of the regression parameter distribution is satisfied. Also, the expression derived is simple, as one only has to define a few matrices. We conduct an extensive Monte Carlo study to illustrate the behavior of the skewness coefficient and we apply it in two real datasets.

Keywords: maximum likelihood estimates; type I and II censoring; skewness coefficient; Weibull censored data

1. Introduction

In its first appearance, the Weibull distribution [1] claimed its wide applicability. Survival analysis, reliability engineering, and extreme value theory are some of its applicability. To amplify the relevance of the Weibull, a regression structure is added to one of the parameters, i.e., the behavior of the distribution may be explained from covariates (explanatory variables) and unknown parameters to be estimated from observable data.

In statistical inference, it is often desirable to test if there are regression parameters statistically significant and the Wald test is commonly performed. Under standard regularity conditions, the null distribution of the Wald statistic is asymptotically chi-squared, a consequence of the maximum likelihood estimators (MLE) distribution. Therefore, the Wald test must be avoided if the sample size is not large enough, because the distribution of the MLE will be poorly approximated by the normal distribution.

Preventing the complexity of the statistical tests, the skewness coefficient (say γ) of the distribution of the MLE is an easy way to verify if the approximation to normality is adequate. A value of γ far from zero indicates a departure from the normal distribution. Pearson's standardized third cumulant defined by $\gamma = \kappa_3/\kappa_2^{3/2}$, where κ_r is the rth cumulant of the distribution, is the most well-known measure of skewness. When $\gamma > 0$ ($\gamma < 0$) the distribution is positively (negatively) skewed and will have a longer (shorter) right tail and a shorter (longer) left tail. If the distribution is symmetrical, γ equals zero. However, there are in [2] (Exercise 3.26) asymmetrical distributions with as many zero-odd order central moments as desired, so, the value of γ must be interpreted with some caution.

In the statistical literature, there is not a closed-form for the skewness coefficient of γ of the MLE in several regression models. Ref. [3] obtained a general $n^{-1/2}$ γ expression (say γ_1) for the distribution of the MLE, where n is the sample size. Following [3], several works have been developed in order to obtain the γ_1 coefficient. In the first, Ref. [4] determined its expression for the class of generalized linear models and, the last one, Ref. [5] defined the γ_1 for the varying dispersion beta regression model and showed that this coefficient for the distribution of the MLE of the precision parameter is relatively large in small to moderate sample sizes. This paper is the first focused on a censored model.

In this work, we derive the γ_1 coefficient of the distribution of the MLE of the linear parameters in the Weibull censored data, assuming σ known, as $\sigma = 1/2$ and 1, the Rayleigh and exponential models, respectively. We discuss the situation when σ is unknown, however, it can be replaced by a consistent estimator, and then we can turn back to the original situation. This type of procedure was performed, for instance, by [4].

The remainder of the paper is organized as follows. Section 2 defines the Weibull censored data. In Section 3, we obtain a simple matrix expression, of order $n^{-1/2}$, for the skewness coefficients of the distributions of the MLEs of the linear regression parameters. In Section 4, some Monte Carlo simulations are performed. Two applications are presented in Section 5. Concluding remarks are offered in Section 6.

2. The Weibull Censored Data

We say that a continuous random variable T has Weibull distribution with scale parameter θ and shape parameter σ, or $T \sim \mathrm{WE}(\theta, \sigma)$, if its probability density function (pdf) is given by

$$f(t;\theta,\sigma) = \frac{1}{\sigma\theta^{1/\sigma}} t^{1/\sigma-1} \exp\left\{-\left(t/\theta\right)^{1/\sigma}\right\}, \tag{1}$$

with $t > 0$, $\sigma > 0$ and $\theta > 0$. From (1), we can observe two particular distributions: the exponential and the Rayleigh, where $\sigma = 1$ and $\sigma = 1/2$, respectively. In lifetime data, there is the censoring restriction, i.e, if $T_1, \ldots, T_n$ are a random sample from (1), instead of T_i, we observe, under right censoring, $t_i = \min(T_i, L_i)$, where L_i is the censoring time, independent of T_i, $i = 1, \ldots, n$. In this work, we consider an hybrid censoring scheme, where the study is finalized when a pre-fixed number, $r \leq n$, out of n observations have failed, as well as when a prefixed time, say $L_1 = \ldots = L_n = L$, has been reached. The type I censoring is a particular case for $r = n$ and the type II censoring appears when $L_1, \ldots, L_n = +\infty$. Additionally, we add the non-informative censoring assumption, i.e., the random variables L_i does not depend on θ. Under this setup, the log-likelihood function has the form

$$L(\theta,\sigma) \quad = \left(\sigma\theta^{1/\sigma}\right)^{-r} \exp\left\{\left(\tfrac{1}{\sigma} - 1\right) A_1 - \tfrac{1}{\theta^{1/\sigma}} A_2\right\},$$

where $r = \sum_{i=1}^n \delta_i$, $A_1 = \sum_{i=1}^n \delta_i \log t_i$, $A_2 = \sum_{i=1}^n t_i^{1/\sigma}$, $\delta_i = 1$, if $T_i \leq L_i$ and $\delta_i = 0$, otherwise. Usually, the regression modeling considers the distribution of $Y_i = \log(T_i)$ instead of T_i. The distribution of Y_i is of the extreme value form with pdf given by

$$f(y_i; \boldsymbol{x}_i) = \frac{1}{\sigma} \exp\left\{\frac{y_i - \mu_i}{\sigma} - \exp\left(\frac{y_i - \mu_i}{\sigma}\right)\right\}, \quad -\infty < y_i < \infty, \tag{2}$$

where $\mu_i = \log\theta_i$. The regression structure can be incorporated in (2) by making $\theta_i = \exp\left(\boldsymbol{x}_i^\top \boldsymbol{\beta}\right)$, where $\boldsymbol{\beta}$ is a p-vector of unknown parameters and $\boldsymbol{x}_i$ is a vector of regressors related to the ith observation. From this moment, we assume that σ is known, then, the log-likelihood function derived from (2) is given by

$$\ell(\boldsymbol{\beta}) = \sum_{i=1}^n \left[\delta_i \left(-n\log\sigma + \frac{y_i - \mu_i}{\sigma}\right) - \exp\left(\frac{y_i - \mu_i}{\sigma}\right)\right].$$

The total score function and the total Fisher information matrix for $\boldsymbol{\beta}$ are, respectively, $\boldsymbol{U}_\beta = \sigma^{-1}\boldsymbol{X}^\top \boldsymbol{W}^{1/2} \boldsymbol{v}$ and $\boldsymbol{K}_{\beta\beta} = \sigma^{-2}\boldsymbol{X}^\top \boldsymbol{W}\boldsymbol{X}$, where $\boldsymbol{X} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)^\top$, the model matrix, assuming rank $(\boldsymbol{X}) = p$, $\boldsymbol{W} = \mathrm{diag}(w_1, \ldots, w_n)$, $w_i = \mathbb{E}\left[\exp\left(\frac{y_i - \mu_i}{\sigma}\right)\right]$ and $\boldsymbol{v} = (v_1, \ldots, v_n)^\top$, $v_i = \left\{-\delta_i + \exp\left(\frac{y_i - \mu_i}{\sigma}\right)\right\} w_i^{-1/2}$. It can observed that the value of w_i depends on the mechanism of censoring. That means $w_i = q \times \left(1 - \exp\left\{-L_i^{1/\sigma} \exp(-\mu_i/\sigma)\right\}\right) + (1-q) \times (r/n)$, where $W_{(r)}$ denotes the rth order statistic from $W_1, \ldots, W_n$ and $q = \mathbb{P}\left(W_{(r)} \leq \log L_i\right)$. Note that $q = 1$ and $q = 0$ for types I and II censoring, respectively. The proof is presented in the Appendix A. The MLE of $\boldsymbol{\beta}$, $\widehat{\boldsymbol{\beta}}$, is the solution of $\boldsymbol{U}_{\boldsymbol{\beta}} = \boldsymbol{0}$. The $\widehat{\boldsymbol{\beta}}$ can not be expressed in closed-form. It is typically obtained by numerically maximizing the log-likelihood function using a Newton or quasi-Newton nonlinear optimization algorithm. Under mild regularity conditions and in large samples,

$$\widehat{\boldsymbol{\beta}} \sim \mathrm{N}_p\left(\boldsymbol{\beta}, \boldsymbol{K}_{\boldsymbol{\beta}\boldsymbol{\beta}}^{-1}\right),$$

approximately.

3. Skewness Coefficient

As discussed, the skewness coefficient is simple way to verify whether the approximation to normality is adequate. The model presented in (2) does not has a closed-form for this coefficient. The alternative is to apply the [3] result. These authors derived an approximation of order $\mathcal{O}(n^{-2})$ for the third cumulant of the MLE of the a-th regressor, i.e.,

$$\kappa_3(\hat{\beta}_a) = \mathbb{E}\left\{\left[\hat{\beta}_a - \mathbb{E}(\hat{\beta}_a)\right]^3\right\},$$

$a = 1, \ldots, p$, which can be expressed as

$$\kappa_3(\hat{\beta}_a) = \sum{}' \kappa^{a,b} \kappa^{a,c} \kappa^{a,d} m_{bc}^{(d)}, \tag{3}$$

where $m_{bc}^{(d)} = 5\kappa_{bc}^{(d)} - (\kappa_{cd}^{(b)} + \kappa_{bd}^{(c)} + \kappa_{bcd})$, $a = 1, \ldots, p$. Here, $\sum'$ represents the summation over all combinations of parameters and over all the observations. From (3), after some algebra, we can express the third cumulant of the distribution of $\hat{\boldsymbol{\beta}}$ for the Weibull censored data as

$$\boldsymbol{\kappa}_3(\widehat{\boldsymbol{\beta}}) = -\sigma^{-3} \boldsymbol{P}^{(3)} \left(\boldsymbol{W} + 3\sigma \boldsymbol{W}'\right) \boldsymbol{1}, \tag{4}$$

where $\boldsymbol{W}' = \mathrm{diag}(w_1', \ldots, w_n')$, $w_i' = -\sigma^{-1} L_i^{1/\sigma} \exp\{-L_i^{1/\sigma} \exp(-\mu_i/\sigma) - \mu_i/\sigma\}$, $\boldsymbol{P} = \boldsymbol{K}_{\boldsymbol{\beta}\boldsymbol{\beta}}^{-1} \boldsymbol{X}^\top = \sigma^2 \left(\boldsymbol{X}^\top \boldsymbol{W} \boldsymbol{X}\right)^{-1} \boldsymbol{X}^\top$, $\boldsymbol{P}^{(3)} = \boldsymbol{P} \odot \boldsymbol{P} \odot \boldsymbol{P}$, $\odot$ represents a direct product of matrices and $\boldsymbol{1}$ is a n-dimensional vector of ones. Finally, by (4) and the Fisher information matrix, the asymmetry coefficient of the distribution of $\widehat{\boldsymbol{\beta}}$ to order $n^{-1/2}$ is given by

$$\gamma_1(\hat{\boldsymbol{\beta}}) = -\sigma^{-3} \boldsymbol{P}^{(3)} \left(\boldsymbol{W} + 3\sigma \boldsymbol{W}'\right) \boldsymbol{1} \odot \left\{\mathrm{diag}\left(\boldsymbol{K}_{\boldsymbol{\beta}\boldsymbol{\beta}}^{-1}\right) \boldsymbol{1}\right\}^{-3/2}, \tag{5}$$

for type II censoring, $\boldsymbol{W}' = \boldsymbol{0}$, then (5) reduces to $\gamma_1(\hat{\boldsymbol{\beta}}) = -\sigma^{-3} \boldsymbol{P}^{(3)} \boldsymbol{W} \boldsymbol{1} \odot \left\{\mathrm{diag}\left(\boldsymbol{K}_{\boldsymbol{\beta}\boldsymbol{\beta}}^{-1}\right) \boldsymbol{1}\right\}^{-3/2}$. More details about the involved expressions are presented in Appendix A. The study of asymptotic properties of the Weibull censored data was the goal of many papers. Refs. [6,7] derived the Bartlett and the Bartlett-type correction factors for likelihood ratio and score tests, respectively, for the exponential censored data. Ref. [8] generalized these previous for the Weibull censored data and also derived the Bartlett-type correction factors for the gradient test. Ref. [9] presented the asymptotic expansions up to order $n^{-1/2}$ of the non null distribution functions of the likelihood ratio, Wald, Rao score and gradient statistics also for the censored exponential data. The result in expression (5) can be incorporated in this gallery.

4. Simulation Study

In this section, we compare the sample skewness coefficient (ρ) and the $n^{-1/2}$ skewness coefficients evaluated in the true and estimated parameters ($\widehat{\gamma}_1^\star$ and $\widehat{\gamma}_1$, respectively) of the distributions of the MLEs in the Weibull censored model. To draw the data, we consider three values for σ: 0.5, 1 and 3; five sample sizes: 20, 30, 40, 60 and 100; three values for the percent of censoring C: 10%, 25% and 50%; and two number of regressors p: 3 and 5, where we consider two vectors for $\boldsymbol{\beta}$ in each case: $(-2, 0.5, 1)$ and $(1, -0.75, 0.5)$ for $p = 3$ and $(-2, 0.5, 1, -0.3, -0.5)$ and $(1, -0.75, 0.5, -1, 0.8)$ for $p = 5$. For each combination of σ, $\boldsymbol{\beta}$, % of censoring and sample size we considered 20,000 Monte Carlo replicates. Each vector of covariates $\boldsymbol{x}_i$ considers an intercept term and the $p - 1$ remaining covariates were drawn independently from the standard normal distribution. Values from the Weibull model are drawn considering the inverse transformation method. Therefore, the greater $n \times C/100$ values were censored at the observed $(1 - C/100)$-th quantile (a type II censoring scheme). For each sample, we considered the jackknife estimator for σ, say $\widehat{\sigma}_J$. Therefore, the computation of $\widehat{\gamma}_1^\star$ and $\widehat{\gamma}_1$ was performed considering $(\boldsymbol{\beta}, \sigma)$ and $(\widehat{\boldsymbol{\beta}}, \widehat{\sigma}_J)$, the true and estimated parameters, respectively. Additionally, ρ is computed based on the 20,000 (marginal) skewness coefficient for the components of $\widehat{\boldsymbol{\beta}}$. Table 1 summarizes the case $\boldsymbol{\beta} = (-2, 0.5, 1)$ (with $p = 3$ regressors) and $C = 10\%$. The main conclusions are the following:

Table 1. The $n^{-1/2}$ and sample skewness coefficients of the distributions of the MLEs in the Weibull censored data with $p = 3$ regressors and $\boldsymbol{\beta} = (-2, 0.5, 1)$.

			$\widehat{\beta}_0$			$\widehat{\beta}_1$			$\widehat{\beta}_2$		
C	σ	n	ρ	$\widehat{\gamma}_1^\star$	$\widehat{\gamma}_1$	ρ	$\widehat{\gamma}_1^\star$	$\widehat{\gamma}_1$	ρ	$\widehat{\gamma}_1^\star$	$\widehat{\gamma}_1$
10%	0.5	20	−0.235	−0.080	−0.095	0.052	0.207	0.197	0.071	0.245	0.234
		30	−0.169	0.009	0.001	0.037	0.040	0.041	0.097	0.212	0.204
		40	−0.114	−0.013	−0.019	0.038	0.107	0.104	0.075	0.196	0.193
		60	−0.139	−0.082	−0.083	0.011	0.061	0.061	0.033	0.171	0.170
		100	−0.059	−0.055	−0.055	0.054	0.074	0.073	−0.005	0.087	0.087
	1.0	20	−0.198	0.017	0.004	−0.008	0.225	0.197	0.068	0.298	0.284
		30	−0.219	−0.101	−0.107	0.110	0.200	0.192	0.216	0.304	0.299
		40	−0.210	0.004	−0.002	0.083	0.140	0.137	0.109	0.185	0.182
		60	−0.147	−0.005	−0.010	0.085	0.143	0.137	0.057	0.173	0.171
		100	−0.092	−0.013	−0.015	0.019	0.048	0.047	0.116	0.132	0.131
	3.0	20	−0.232	0.006	0.005	0.094	0.143	0.131	0.022	0.093	0.087
		30	−0.178	−0.041	−0.040	0.004	0.054	0.048	−0.007	0.147	0.136
		40	−0.185	0.005	0.003	0.002	0.024	0.023	0.064	0.156	0.151
		60	−0.128	−0.020	−0.020	0.044	0.049	0.047	0.073	0.124	0.120
		100	−0.117	−0.028	−0.028	0.084	0.100	0.095	0.039	0.077	0.075

- The terms $\widehat{\gamma}_1^\star$ and $\widehat{\gamma}_1$ are closer in all the considered combinations, suggesting that $\widehat{\gamma}_1$ approaches $\widehat{\gamma}_1^\star$ in a reasonable way, even when the sample size is small.
- In general terms, $\widehat{\gamma}_1$ approaches well ρ for $\widehat{\beta}_1$ and $\widehat{\beta}_2$. However, for $\widehat{\beta}_0$ the terms seem discrepant even for $n = 100$.
- Considering the 90 cases for $p = 3$, ρ ranges from $(-0.245, 0.255)$, $(-0.429, 0.340)$ and $(-0.819, 1.181)$ for C 10%, 25% and 50%, respectively. For $p = 5$, ρ ranges from $(-0.373, 0.252)$, $(-0.402, 0.198)$ and $(-0.787, 0.495)$ for C 10%, 25% and 50%, respectively. This suggest that a higher percentage of censorship produce a higher skewness in the MLE estimators for the components of $\boldsymbol{\beta}$.
- Considering the 90 cases for $p = 3$, ρ ranges from $(-0.819, 1.181)$, $(-0.363, 0.867)$, $(-0.351, 0.411)$, $(-0.305, 0.346)$ and $(-0.273, 0.255)$, for $n = 20, 30, 40, 60$ and 100, respectively. For $p = 5$, ρ ranges

from $(-1.015, 0.740)$, $(-0.529, 0.426)$, $(-0.372, 0.413)$, $(-0.318, 0.320)$ and $(-0.225, 0.243)$ for $n = 20, 30, 40, 60$ and 100, respectively. This suggest that, as expected, when n increases the skewness coefficient of the MLE estimators for the components of $\boldsymbol{\beta}$ will be more symmetric.

Results suggest that, even with a moderate percentage of censored observations and small sample sizes, the distribution of the MLE for the components of $\boldsymbol{\beta}$ in the Weibull censored model are closer to the symmetry. The combinations of $\boldsymbol{\beta}$, p and C not seem to affect the results. A simulation study showing this finding was omitted for the sake of brevity.

5. Applications

In this section we illustrate with two real dataset the application of the estimated skewness coefficient for the MLE estimators in the Weibull censored regression model. All the routine was performed in the statistical software R, [10]. Codes can be found in the personal website from the first author https://www.ufjf.br/tiago_magalhaes/downloads/.

5.1. Smokers Dataset

This dataset is related to a clinical trial on the effectiveness of triple-combination pharmacotherapy for tobacco dependence treatment conducted by the Cancer Institute of New Jersey and Robert Wood Johnson Foundation. The trial recruited 127 smokers 18 years or older with predefined medical illnesses from the local community. The outcome were the `time` (in days) to first relapse (return to smoking). The study lasted 182 days (26 weeks). Therefore, the times are subject to a censoring type I (32% of times were censored). We only considered the 113 patients where such observed time was positive (non-zero). Other measures were assigned randomly treatment group with levels combination or patch only (`grp`), age in years at time of randomization (`age`) and `employment` (full-time or non-full-time). We consider that $\texttt{time}_i \sim \text{WE}(\theta_i; \sigma)$, where $\log \theta_i = \boldsymbol{X}_i^\top \boldsymbol{\beta}$, $\boldsymbol{\beta} = (\beta_{\texttt{intercept}}, \beta_{\texttt{grp}}, \beta_{\texttt{age}}, \beta_{\texttt{employment}})^\top$ and

$$\boldsymbol{X}_i^\top = (1, \texttt{grp}_i, \texttt{age}_i, \texttt{employment}_i)$$

We estimated $\widehat{\sigma}_J = 1.617008$ based on the jackknife method, which was used as known in all the computations. Table 2 shows the parameters estimates, their standard errors and the estimated skewness coefficients and Figure 1 shows the estimated density function based on 1000 bootstrap samples for the coefficients related to the covariates `grp`, `age` and `employment`. Note that the estimated skewness for all parameters were closer to zero, suggesting a symmetric distribution for the estimators which is corroborated by the estimated density based on the bootstrap.

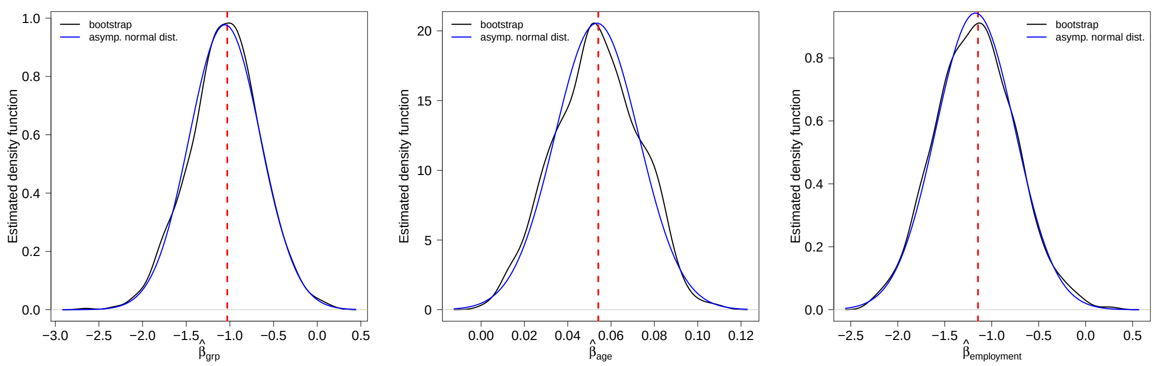

Figure 1. Estimated density function based on 1000 bootstrap samples and the asymptotic distribution for $\widehat{\beta}_{\texttt{grp}}$ (left panel), $\widehat{\beta}_{\texttt{age}}$ (center panel) and $\widehat{\beta}_{\texttt{employment}}$ (right panel). The red line denotes the estimated parameter.

Table 2. Estimates for parameters and skewness coefficient in smokers dataset.

Parameter	Estimate	s.e.	γ_1
$\beta_{\texttt{intercept}}$	3.1690	0.8136	−0.0478
$\beta_{\texttt{grp}}$	−1.0303	0.3694	−0.0529
$\beta_{\texttt{age}}$	0.0541	0.0167	0.1251
$\beta_{\texttt{employment}}$	−1.1460	0.3935	−0.0753

5.2. Insulating Fluids Dataset

This dataset was presented in [11] on insulating fluids and it is related an accelerated test performed in order to determine the relationship between `time` (in minutes) to breakdown and `voltage` (in kilovolts). The authors assumed a regression structure based on the Weibull model and a common censoring time at $L = 200$ (type I censoring), i.e., $\texttt{time}_i \sim \text{WEI}(\theta_i, \sigma)$, where $\log \theta_i = \boldsymbol{X}_i^\top \boldsymbol{\beta}$, $i = 1, \ldots, 76$, $\mathbf{X}_i^\top = (\beta_{\texttt{Intercept}}, \beta_{\texttt{log-voltage}})$. We estimated $\widehat{\sigma}_J = 1.296704$ based on the jackknife method, which was used as known. Table 3 shows the estimates, standard errors and estimated skewness coefficient for the MLE estimators and Figure 2 shows the estimated density function for $\widehat{\beta}_{\texttt{Intercept}}$ and $\widehat{\beta}_{\texttt{log-voltage}}$. Newly, the estimated skewness for both parameters are closer to zero, suggesting a symmetric distribution for the estimators as also suggest the estimated density based on bootstrap.

Table 3. Estimates for parameters and skewness coefficient in insulating fluids dataset.

Parameter	Estimate	s.e.	γ_1
$\beta_{\texttt{intercept}}$	20.4342	1.8772	0.1451
$\beta_{\texttt{log-voltage}}$	−0.5311	0.0557	−0.1517

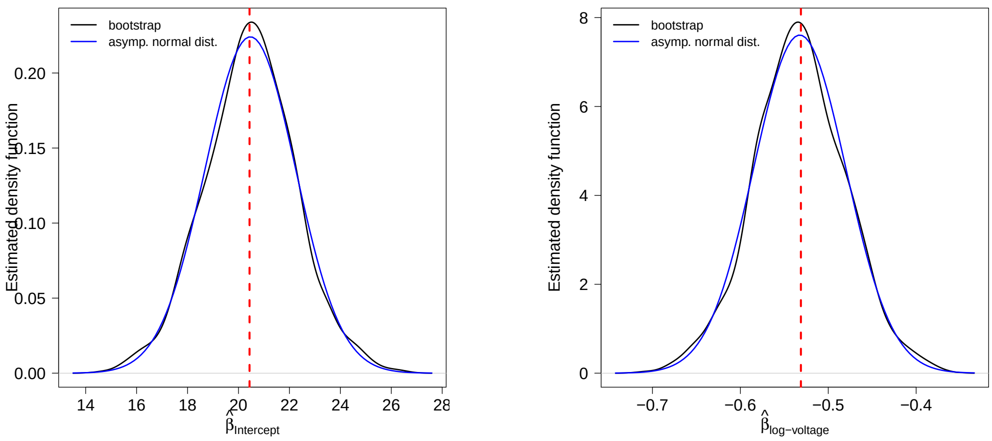

Figure 2. Estimated density function based on 1000 bootstrap samples and the asymptotic distribution for $\widehat{\beta}_{\texttt{Intercept}}$ (left panel) and $\widehat{\beta}_{\texttt{log-voltage}}$ (right panel). The red line denotes the estimated parameter.

6. Concluding Remarks

Since its beginning, the Weibull distribution and regression showed how it is important. In the frequentist context, this model depends strongly on the asymptotic properties of the MLE. Here, we presented an expression of the skewness that, in practical applications, can be used as an indicator of departure from the normal distribution of the MLE. Although the expression (3) entails a great deal of algebra, the final expression (5) of the skewness of the MLE distribution has a very nice form only involving simple operations on diagonal matrices and can be easily implemented into a statistical software, for instance, R, [10].

Author Contributions: All the authors contributed significantly to the present paper.

Funding: The research of D.I. Gallardo was supported by Grant from FONDECYT (Chile) 11160670. H.W. Gómez was supported by Grant SEMILLERO UA-2019 (Chile).

Acknowledgments: We also acknowledge the referee's suggestions that helped us to improve this work.

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

Appendix A

In this Section we provided some additional details related to the computation of the manuscript.

Appendix A.1. W's Quantities

In order to compute $w_i = \mathbb{E}\left[\exp\left(\frac{y_i-\mu_i}{\sigma}\right)\right]$, $i = 1, \ldots, n$, we first study the case type I censoring. Note that

$$\exp\left(\frac{y_i - \mu_i}{\sigma}\right) = \begin{cases} \exp\left(\frac{y_i-\mu_i}{\sigma}\right), & \text{if } y_i \leq \log L_i \\ \exp\left(\frac{\log L_i-\mu_i}{\sigma}\right), & \text{otherwise} \end{cases}$$

Therefore,

$$\begin{aligned} w_i &= \int_{-\infty}^{\log L_i} \frac{1}{\sigma} \exp\left(\frac{2(y_i-\mu_i)}{\sigma} - \exp\left(\frac{y_i-\mu_i}{\sigma}\right)\right) dy_i + \exp\left(\frac{\log L_i - \mu_i}{\sigma}\right) \mathbb{P}(T_i > L_i) \\ &= 1 - \exp\left(-L_i^{1/\sigma} e^{-\mu_i/\sigma}\right)\left(1 + L_i^{1/\sigma} e^{-\mu_i/\sigma}\right) + L_i^{1/\sigma} e^{-\mu_i/\sigma} \exp\left(-L_i^{1/\sigma} e^{-\mu_i/\sigma}\right) \\ &= 1 - \exp\left(-L_i^{1/\sigma} e^{-\mu_i/\sigma}\right). \end{aligned}$$

Direct computation also shows that

$$\begin{aligned} v_i &= \mathbb{E}\left[\exp\left(\frac{2(y_i-\mu_i)}{\sigma}\right)\right] \\ &= \int_{-\infty}^{\log L_i} \frac{1}{\sigma} \exp\left(\frac{3(y_i-\mu_i)}{\sigma} - \exp\left(\frac{y_i-\mu_i}{\sigma}\right)\right) dy_i + \exp\left(\frac{2(\log L_i - \mu_i)}{\sigma}\right) \mathbb{P}(T_i > L_i) \\ &= 2 - \exp\left(-L_i^{1/\sigma} e^{-\mu_i/\sigma}\right)\left[2 + 2L_i^{1/\sigma} e^{-\mu_i/\sigma} + L_i^{2/\sigma} e^{-2\mu_i/\sigma}\right] + L_i^{2/\sigma} e^{-2\mu_i/\sigma} \exp\left(-L_i^{1/\sigma} e^{-\mu_i/\sigma}\right) \\ &= 2\left\{1 - \exp\left(-L_i^{1/\sigma} e^{-\mu_i/\sigma}\right)\left[1 + L_i^{1/\sigma} e^{-\mu_i/\sigma}\right]\right\} \\ &= 2\left\{w_i + \sigma w_i'\right\}. \end{aligned}$$

On the other hand, for the type II censoring note that $W_i = \exp\left(\frac{y_i-\mu_i}{\sigma}\right) \sim E(1)$. By [12], we have that

$$W_{(i)} \overset{\mathcal{D}}{=} \sum_{j=1}^{i} \frac{Z_j}{n-j+1}, \tag{A1}$$

where $W_{(i)}$ is the ith order statistic from $W_1, \ldots, W_n$, $\overset{\mathcal{D}}{=}$ denotes "equal in distribution" and $Z_1, \ldots, Z_n$ are independent and identically distributed $E(1)$ random variables. Therefore, $\mathbb{E}(W_{(j)}) = \mathrm{Var}(W_{(j)}) = \sum_{k=1}^{j}(n-k+1)^{-1}$ and

$$W_i = \begin{cases} W_{(1)}, & \text{with probability } 1/n \\ \vdots & \\ W_{(r-1)}, & \text{with probability } 1/n \\ W_{(r)}, & \text{with probability } (n-r+1)/n \end{cases}$$

Therefore,

$$w_i = \mathbb{E}(W_i) = \frac{1}{n}\sum_{j=1}^{r-1}\sum_{k=1}^{j}(n-k+1)^{-1} + \frac{(n-r+1)}{n}\sum_{k=1}^{r}(n-k+1)^{-1}.$$

With some manipulations, we obtain that $\mathbb{E}(W_i) = r/n$. Also, we have that $\mathbb{E}(W_{(j)}^2) = \mathbb{E}(W_{(j)}) + \mathbb{E}^2(W_{(j)})$ and

$$V_i = \begin{cases} W_{(1)}^2, & \text{with probability } 1/n \\ \vdots & \\ W_{(r-1)}^2, & \text{with probability } 1/n \\ W_{(r)}^2, & \text{with probability } (n-r+1)/n \end{cases}$$

Therefore,

$$v_i = w_i + \frac{1}{n}\sum_{j=1}^{r-1}\left[\sum_{k=1}^{j}(n-k+1)^{-1}\right]^2 + \frac{(n-r+1)}{n}\left[\sum_{k=1}^{r}(n-k+1)^{-1}\right]^2.$$

Algebraic manipulations shows that

$$v_i = \frac{1}{n}\left[r + \sum_{k=1}^{r}\frac{2(r-k)+1}{n-k+1}\right].$$

Finally, as the hybrid scheme can be seen as a mixture between type I and II censoring, we obtain directly that

$$\begin{aligned} w_i &= q \times \left(1 - \exp\left(-L_i^{1/\sigma}e^{-\mu_i/\sigma}\right)\right) + (1-q) \times (r/n), \\ v_i &= q \times 2\left\{1 - \exp\left(-L_i^{1/\sigma}e^{-\mu_i/\sigma}\right)\left[1 + L_i^{1/\sigma}e^{-\mu_i/\sigma}\right]\right\} + (1-q) \times \frac{1}{n}\left[r + \sum_{k=1}^{r}\frac{2(r-k)+1}{n-k+1}\right], \end{aligned}$$

where q is the mixing probability given by $q = \mathbb{P}(W_{(r)} \leq \log L)$. By (A1), $W_{(r)}$ has hypoexponential [13] distribution with vector of parameters $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n)$, where $\lambda_j = (n-j+1)^{-1}$. Therefore,

$$q = 1 - \sum_{j=1}^{n}\frac{L^{-\lambda_j}}{P_j},$$

where $P_j = \prod_{k=1, k\neq j}^{n}(k-j)/(n-j+1)$.

Appendix A.2. Derivatives and Cumulants

Let $Y_1, \ldots, Y_n$ a random sample from Weibull censored data, the logarithm of the likelihood function is given by

$$\ell(\boldsymbol{\beta}) = \sum_{i=1}^{n} \left\{ \delta_i \left[-n \log \sigma + \frac{y_i - \mu_i}{\sigma} \right] - \exp\left(\frac{y_i - \mu_i}{\sigma} \right) \right\}. \tag{A2}$$

The first four derivatives of (A2) can be expressed, respectively, for

$$\frac{\partial}{\partial \beta_r} \ell(\boldsymbol{\beta}) = \frac{1}{\sigma} \sum_{i=1}^{n} \left\{ -\delta_i + \exp\left(\frac{y_i - \mu_i}{\sigma} \right) \right\} x_{ri};$$
$$\frac{\partial^2}{\partial \beta_r \partial \beta_s} \ell(\boldsymbol{\beta}) = -\frac{1}{\sigma^2} \sum_{i=1}^{n} \exp\left(\frac{y_i - \mu_i}{\sigma} \right) x_{ri} x_{si};$$
$$\frac{\partial^3}{\partial \beta_r \partial \beta_s \partial \beta_t} \ell(\boldsymbol{\beta}) = \frac{1}{\sigma^3} \sum_{i=1}^{n} \exp\left(\frac{y_i - \mu_i}{\sigma} \right) x_{ri} x_{si} x_{ti}.$$

The second- to third-order cumulants are

$$\kappa_{rs} = -\frac{1}{\sigma^2} \sum_{i=1}^{n} w_i x_{ri} x_{si};\ \kappa_{r,s} = -\kappa_{rs} = \frac{1}{\sigma^2} \sum_{i=1}^{n} w_i x_{ri} x_{si};$$

$$\kappa_{rst} = \frac{1}{\sigma^3} \sum_{i=1}^{n} w_i x_{ri} x_{si} x_{ti};\ \kappa_{rs}^{(t)} = -\frac{1}{\sigma^2} \sum_{i=1}^{n} w_i' x_{ri} x_{si} x_{ti};$$

where $w_i = \mathbb{E}\left\{ \exp\left(\frac{y_i - \mu_i}{\sigma} \right) \right\}$,

$$w_i = 1 - \exp\left\{ -L_i^{1/\sigma} \exp(-\mu_i/\sigma) \right\},$$
$$w_i' = -\frac{1}{\sigma} L_i^{1/\sigma} \exp\{ -L_i^{1/\sigma} \exp(-\mu_i/\sigma) - \mu_i/\sigma \},$$

It can be observed that $w_i' = 0$ for type II censoring.

References

1. Weibull, W. A statistical distribution function of wide applicability. *J. Appl. Mech.* **1951**, *18*, 293–297.
2. Kendall, M.G.; Stuart, A. *The Advanced Theory of Statistic*, 4th ed.; Griffin: London, UK, 1977.
3. Bowman, K.O.; Shenton, L.R. Asymptotic skewness and the distribution of maximum likelihood estimators. *Commun. Stat. Theory Methods* **1998**, *27*, 2743–2760.
4. Cordeiro, H.H.; Cordeiro, G.M. Skewness for parameters in generalized linear models. *Commun. Stat. Theory Methods* **2001**, *30*, 1317–1334.
5. Magalhães, T.M.; Botter, D.A.; Sandoval, M.C.; Pereira, G.H.A.; Cordeiro, G.M. Skewness of maximum likelihood estimators in the varying dispersion beta regression model. *Commun. Stat. Theory Methods* **2019**, *48*, 4250–4260.
6. Cordeiro, G.M.; Colosimo, E.A. Improved likelihood ratio tests for exponential censored data. *J. Stat. Comput. Sim.* **1997**, *56*, 303–315.
7. Cordeiro, G.M.; Colosimo, E.A. Corrected score tests for exponential censored data. *Stat. Probab. Lett.* **1999**, *44*, 365–373.
8. Magalhães, T.M.; Gallardo, D.I. Bartlett and Bartlett-type corrections for Weibull censored data. *Stat. Oper. Res. Trans.* **2019**, Submitted.

9. Lemonte, A.J. Non null asymptotic distribution of some criteria in censored exponential regression. *Commun. Stat. Theory Methods* **2014**, *43*, 3314–3328.
10. R Development Core Team. *A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2019.
11. Nelson, W.; Meeker, W.Q. Theory for optimum accelerated censored life tests for Weibull and extreme value distributions. *Technometrics* **1978**, *20*, 171–177.
12. Renyi, A. On the theory of order statistics. *Acta Math. Hung.* **1953**, *4*, 191–231.
13. Devianto, D.; Oktasari, L.; Maiyastri. Some properties of hypoexponential distribution with stabilizer constant. *Appl. Math. Sci.* **2015**, *142*, 7063–7070.