Gulf Journal of Mathematics ISSN: 2309-4966
Vol 23, Issue 1 (2026) 1-20 https://doi.org/10.56947/jv413h17

WEIBULL TAIL ESTIMATION VIA PEAK-OVER-THRESHOLD UNDER DETERMINISTIC THRESHOLDS

NESRINE IDIOU[1], MOHAMMED RIDHA KOUIDER[2] and FATAH BENATIA[3*]

ABSTRACT. This paper investigates the estimation of the Weibull tail coefficient within the framework of extreme value theory, using the Peak-Over-Threshold approach with a deterministic threshold. Based on maximum likelihood estimation, the methodology offers a unified treatment of the shape parameter and the finite right endpoint, while addressing a key challenge in modeling bounded extremes. Theoretical properties such as existence, consistency, and asymptotic normality of the estimators are established. Furthermore, a new log-excess-based estimator for the Weibull tail coefficient is introduced. The performance of the new estimators in small samples is illustrated through simulations, and their effectiveness, robustness, and practical relevance are further confirmed by real-data applications. This dual validation highlights both their solid theoretical foundation and applied significance.
Keywords. Extreme value index, Weibull distribution, Maximum likelihood estimation, Peak-Over-Threshold, Generalized Pareto distribution.
2020 Mathematics Subject Classification. Primary 62G32; Secondary 60G70.

1. INTRODUCTION

Assume that $X_1, X_2, \ldots, X_n$ are independent and identically distributed (i.i.d) random variables with common distribution function F. We define the upper bound of the support F as $\tau_F := \sup\{x : F(x) < 1\} < \infty$. Furthermore, for $x > 0$, the survival function associated with F is expressed as $\overline{F}(x) = 1 - F(x)$, and $F(x) := \int_0^x dF(y)$ is the distribution function. The empirical counterparts based on the sample $X_1, X_2, \ldots, X_n$ are given by

$$F_n(x) := \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\{X_i \le x\}}, \quad \overline{F}_n(x) := \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\{X_i > x\}}. \tag{1}$$

Noted $\overline{F}(x) = 1 - F(x)$, which corresponds to the complementary cumulative distribution function of X. Within the framework of extreme value theory (EVT), the tail behavior of $\overline{F}(x)$ is characterized by a parameter known as the extreme value index, denoted by γ for the distribution of X which controls how the survival function decays.

Date: Received: Mar 14, 2026; Revised: Apr 11, 2026; Accepted: Apr 25, 2026.
* Corresponding author
© The Author(s) 2025. This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License. To view a copy of the licence, visit https://creativecommons.org/licenses/by-nc-nd/4.0/.

If $\gamma > 0$, then $\overline{F}$ is regarded as heavy-tailed. If $\gamma < 0$, then $\overline{F}$ vanishes above its finite right endpoint $\tau_F < \infty$. If $\gamma = 0$, then $\overline{F}$ decreases to zero at an exponential rate. Given that $\overline{F}$ has rank $-1/\gamma$ and is frequently shifting at infinite levels, we write $\overline{F} \in RV_{-1/\gamma}$. Therefore, for each $x > 0$,

$$\lim_{t\to\infty} \frac{\overline{F}(tx)}{\overline{F}(t)} = x^{-1/\gamma}. \tag{2}$$

In this work, we assume that F is of Pareto type, where τ_F represents an appropriate upper bound of the support of F and $\gamma \in \mathbb{R}$. The distribution falls into one of the following three cases:

Case 1: $\gamma > 0$. In this case, $\overline{F}(x) = x^{-1/\gamma}\,\ell_F(x)$, here ℓ_F is a function that grows progressively at infinity.

Case 2: $\gamma < 0$. In this case, $\overline{F}(x) = (\tau_F - x)^{-1/\gamma}\,\ell_F\left(\frac{1}{\tau_F - x}\right)$, where $\tau_F < \infty$ the finite right endpoint.

Case 3: $\gamma = 0$. In this case, $\overline{F}(x) = e^{-x/\sigma}$, where σ denotes a positive scale parameter. Finally, ℓ_F denotes a slowly varying function at infinity i.e., $\forall x > 0, \lim_{t\to\infty} \dfrac{\ell_F(tx)}{\ell_F(t)} = 1$. We apply maximum likelihood estimation (MLE) in this work, considering that the entire sample is taken from an extreme value distribution. As a representative example, we consider the Weibull distribution, whose tails exhibit an exponentially fast decay. For $\gamma < 0$, the Weibull law admits the following cumulative distribution function (cdf):

$$G_{\gamma<0}(x) = \begin{cases} e^{-(-x)^{-1/\gamma}} & \text{if } x \le 0 \\ 1 & \text{if } x > 0 \end{cases} \tag{3}$$

In addition, F falls into the Weibull region of attraction precisely when

$$\tau_F := \sup\{x \in \mathbb{R} : F(x) < 1\} < \infty,$$

with the tail distribution fulfilling

$$\overline{F}(\tau_F - x) \in RV_{-1/\gamma}, \quad \text{for } \gamma < 0.$$

These conditions are firmly established in the classical literature, particularly in Balkema & de Haan (1974)[1] and in the monograph of de Haan & Ferreira (2006)[7] on extreme value theory, where they provide a rigorous characterization in the Weibull attraction class.

The Weibull model is a prominent lifetime model, widely applied in medicine, biology, and engineering sciences, and has been extensively studied over the past three decades [5]. Among the various estimation procedures proposed for the Weibull index, maximum likelihood estimation remains the most classical and widely implemented, well known for its consistency and convergence. In this study, we focus on the maximum likelihood estimator of the generalized Pareto distribution (GPD) within the Peak-Over-Threshold framework, under a deterministic threshold t and for $\gamma \in [-1, 0)$. Take $X_1, X_2, \ldots, X_n$ to be i.i.d according to F, and write $X_{1,n} < \cdots < X_{n,n}$ for the sample arranged in increasing order, known as the order statistics.

In particular, the i-th order statistic satisfies $X_{i,n} = X_j$ for some $1 \leq j \leq n$. The maximum likelihood estimation procedure builds upon the seminal contributions of Balkema and de Haan [1] and Pickands [22], who demonstrated that the limiting behavior of the exceedance distribution

$$C_j = X_j - t, \quad \text{conditional on } X_j > t,$$

as the threshold t approaches τ_F, the exceedance distribution converges to the GPD. The amount of absolute exceedances beyond the threshold t, denoted by k, corresponds to the greatest values $(X_{n-k,n}, \ldots, X_{n,n})$. Next, applying maximum likelihood inside the GPD framework, the unidentified components γ and σ are estimated by maximizing

$$\ell(C_j; \gamma, \sigma) = \prod_{j=1}^{k} f_{\gamma,\sigma}(C_j), \tag{4}$$

where $f_{\gamma,\sigma}(x) = \frac{1}{\sigma}\left(1 + \frac{\gamma}{\sigma}x\right)^{-1/\gamma-1}$, denotes the density function of the GPD, and $C_j = X_j - t$ for $1 \leq j \leq k$. The log-likelihood formula that corresponds to this is provided by

$$\log \ell(C_j; \gamma, \sigma) = \sum_{j=1}^{k} \left[\log\left(\frac{1}{\sigma}\right) - \left(\frac{1}{\gamma} + 1\right) \log\left(1 + \frac{\gamma}{\sigma} C_j\right)\right]. \tag{5}$$

A scale parameter allowable range σ is given by $\sigma > 0$ when $\gamma > 0$, even $\sigma > -\gamma C_{k,k}$ when $\gamma \leq 0$. In particular, if $\gamma < -1$, the MLE of the GPD parameters does not exist. We confine attention to the parameter space, as our goal is to derive the Weibull shape parameter $A = \{\sigma > -\gamma C_{k,k}; -1 \leq \gamma < 0\}$. Where $C_{k,k} = \max\{C_j : 1 \leq j \leq k\}$ denotes the largest exceedance. The likelihood equations derived from (5) in terms of the exceedances C_j, for $\gamma \neq 0$, are expressed as

$$\begin{cases} \frac{1}{k}\sum\limits_{j=1}^{n} (1 - \theta C_j)^{-1} 1_{X_j > t} = \frac{1}{\gamma+1} \\ \frac{1}{k}\sum\limits_{j=1}^{n} \log(1 - \theta C_j) 1_{X_j > t} = \gamma, \end{cases} \tag{6}$$

with $\gamma \geq -1$ and $\theta = -(\gamma/\sigma)$. In addition, Kouider et al. [14, 15] elaborated on assessing the extreme value index through a GPD parameters [21]. Their analysis of the profile log-likelihood function revealed that it admits two solutions, denoted by $\left(\widehat{\gamma}^{KIB}, \widehat{\sigma}^{KIB}\right)$. Building on this result, we propose a procedure for estimating $\widehat{\gamma}_t^{KIB}$ and $\widehat{\sigma}_t^{KIB}$ within the Peak-Over-Threshold framework, under a deterministic threshold t, as follows:

We first obtain the underlying estimator $\widehat{\theta}_t^{KIB}$ as the solution of the equation $\psi(\theta) = 0$, where

$$\psi(\theta) = \left(\frac{1}{k}\sum_{j=1}^{n} \log(1 - \theta C_j) 1_{\{X_j > t\}} + 1\right) \frac{1}{k}\sum_{j=1}^{n} (1 - \theta C_j)^{-1} 1_{\{X_j > t\}} - 1. \tag{7}$$

Once $\widehat{\theta}_t^{KIB}$ is obtained, we compute the estimators $\widehat{\gamma}_t^{KIB}$ and $\widehat{\sigma}_t^{KIB}$ respectively as

$$\widehat{\gamma}_t^{KIB} = \frac{1}{k}\sum_{j=1}^{n} \log\left(1-\widehat{\theta}_t^{KIB} C_j\right) 1_{\{X_j>t\}}, \tag{8}$$

and

$$\widehat{\sigma}_t^{KIB} = -\frac{\widehat{\gamma}_t^{KIB}}{\widehat{\theta}_t^{KIB}}.$$

The numerical aspects of solving these equations were investigated by Kouider and Benatia (2023) [13], under the condition $\gamma \geq -1$. In contrast, the following system should be regarded as the empirical counterpart of equation (6), and is expressed as

$$\begin{cases} \dfrac{1}{k}\displaystyle\sum_{j=1}^{n} \left(1-(1-\theta C_j)^{-1}\right) 1_{\{X_j>t\}} = \dfrac{\widehat{\gamma}_t^{KIB}}{\widehat{\gamma}_t^{KIB}+1} \\ \dfrac{1}{k}\displaystyle\sum_{j=1}^{n} \log(1-\theta C_j)\, 1_{\{X_j>t\}} = \widehat{\gamma}_t^{KIB}. \end{cases} \tag{9}$$

We now define the auxiliary function

$$\widetilde{\psi}(\theta) = \left(\frac{1}{k}\sum_{j=1}^{n} \log(1-\theta C_j)\, 1_{\{X_j>t\}} + 1\right) \frac{1}{k}\sum_{j=1}^{n} \left(1-(1-\theta C_j)^{-1}\right) 1_{\{X_j>t\}} + \frac{1}{k}\sum_{j=1}^{n} \log(1-\theta C_j)\, 1_{\{X_j>t\}}.$$

The estimator $\widehat{\theta}_t^{KIB}$ is then obtained as the solution of $\widetilde{\psi}(\theta) = 0$.

We reformulate the system of maximum likelihood equations in terms of the empirical survival function $\overline{F}_n(x)$ in the analysis that follows. This representation provides a natural framework for deriving an explicit expression of the EVI estimator $\widehat{\gamma}_t^{KIB}$. It further enables the application of the mathematical technique introduced by Beirlant et al. [2], valid for $\widehat{\gamma} \in [-1, 0)$, originally developed in the context of censored data and here adapted to complete data, thereby extending its applicability.

The rest of the article is organized as outlined below. The second part presents a new Weibull tail coefficient estimator, introduced under a deterministic threshold, together with its existence, consistency, and asymptotic properties. Next, the approximate normality of the KIB estimators is rigorously demonstrated in Section 3. A controlled simulation analysis and a proposal to the real world data of Chin et al. (2006)[20], with a focus on the estimation of the Weibull tail coefficient, are carried out in Section 4 to illustrate the efficacy of the suggested methodology. Concluding remarks are provided in Section 5. The appendix in Section 6 presents a lemma that underpins the approximate normalization proof, reinforcing the theoretical basis of the findings.

2. Existence, consistency, and asymptotic properties of KIB estimators

The asymptotic properties of the MLE have been extensively studied in the literature. Since no closed-form expression exists to distinguish it from other estimation methods, the MLE is defined implicitly as an ensemble of problems's resolution (6). Its theoretical properties must therefore be established through alternative approaches.

Drees et al. [9] proved the asymptotic normality of any solution to (6) for the Weibull index (or equivalently, the shape parameter of the Weibull distribution) under the follow-up setting, provided that $-\frac{1}{2} < \gamma < 0$, the threshold is chosen as $t = X_{n-k,n}$. In contrast, Chen and Zhou [25] established the validity of the estimator for $-1 < \gamma < 0$ subject to the domain of attraction assumption (first-order), without requiring the second-degree assumption, again with $t = X_{n-k,n}$.

More recently, Kouider et al. [14, 15] demonstrated that the KIB estimators coincide with the MLE of the GPD parameters obtained when the entire sample is considered, and that they exhibit reduced bias compared to the estimators of Drees et al. [9] for the Weibull index with $t = X_{n-k,n}$. Furthermore, Kouider et al. [14] have shown that when the threshold is random, the approximate normal behavior, associated with the KIB estimators depends upon the GPD parameters, i.e., $t = X_{n-k,n}$. This conclusion is consistent with the framework developed by de Haan and Ferreira [7]. Recall that the Weibull index is the central focus of this study when $-1 \leq \gamma < 0$. Our objective is to establish the existence, consistency, and approximate normality of the estimator under a deterministic threshold t. To this end, we adapt the two systems of equations (6) and (9) as follows. We define the functions

$$\varphi_1(x) = 1 - (1-\theta x)^{-1}, \qquad \varphi_2(x) = \log(1-\theta x), \tag{10}$$

with $\theta = -\gamma/\sigma$. Then, for $s = 1, 2$, we obtain

$$\frac{n}{k}\int_t^{\tau_F} \varphi_s(x-t)\, dF_n(x) = \frac{1}{k}\sum_{j=1}^{n} \varphi_s(X_j - t)\, \mathbf{1}_{\{X_j > t\}},$$

where F_n is the empirical distribution defined in (1). Consequently,

$$\frac{n}{k}\int_t^{\tau_F} \varphi_s(x-t)\, dF_n(x) = -\frac{n}{k}\int_t^{\tau_F} \varphi_s(x-t)\, d\overline{F}_n(x),$$

provided that $\overline{F}_n$ represents the empirical survival law given in (1). Integration by parts yields

$$-\frac{n}{k}\int_t^{\tau_F} \varphi_s(x-t)\, d\overline{F}_n(x) = -\frac{n}{k}\Big(\overline{F}_n(\tau_F)\varphi_s(\tau_F - t) - \overline{F}_n(t)\varphi_s(0) - \int_t^{\tau_F} \overline{F}_n(x)\, d\varphi_s(x-t)\Big).$$

Since $\overline{F}_n(\tau_F) \to 0$ as $t \to \tau_F$, and noting that $\varphi_1(0) = 0$, $\varphi_2(0) = 0$, we deduce

$$\frac{n}{k}\int_t^{\tau_F} \varphi_s(x-t)\, dF_n(x) = \frac{1}{k}\sum_{j=1}^{n} \varphi_s(X_j - t)\, \mathbf{1}_{\{X_j > t\}} = \int_t^{\tau_F} \overline{F}_n(x)\, d\varphi_s(x-t), \quad s = 1, 2. \tag{11}$$

By definition, k denotes the number of exceedances above the threshold t. It follows that k is distributed according to a binomial law $k \sim B\big(n, \overline{F}(t)\big)$. Furthermore, using the rule of large numbers, we arrive at $\overline{F}_n(t) = \frac{k}{n} \xrightarrow{\mathbb{P}} \overline{F}(t)$, as $t \to \tau_F$. Substituting this relation into (9), the system can be rewritten as

$$\begin{cases} \frac{1}{\overline{F}_n(t)} \int\limits_t^{\tau_F} \overline{F}_n(x)\, d\varphi_1(x-t) = \frac{\widehat{\gamma}_t^{KIB}}{\widehat{\gamma}_t^{KIB}+1} \\ \frac{1}{\overline{F}_n(t)} \int\limits_t^{\tau_F} \overline{F}_n(x)\, d\varphi_2(x-t) = \widehat{\gamma}_t^{KIB} \end{cases} \tag{12}$$

For convenience under (12), we set

$$\begin{cases} \frac{1}{\overline{F}(t)} \int\limits_t^{\tau_F} \overline{F}(x)\, d\varphi_1(x-t) = \frac{\gamma}{\gamma+1} \\ \frac{1}{\overline{F}(t)} \int\limits_t^{\tau_F} \overline{F}(x)\, d\varphi_2(x-t) = \gamma \end{cases}$$

It is also possible to recover the system of equations (6). Indeed, we have

$$\frac{1}{k}\sum_{j=1}^{n} (1-\theta C_j)^{-1}\, \mathbf{1}_{\{X_j>t\}} = -\frac{1}{\overline{F}_n(t)} \int_t^{\tau_F} \varphi_3(x-t)\, d\overline{F}_n(x),$$

if $\varphi_3(x-t) = (1-\theta(x-t))^{-1}$. Hence, the system of equations (6) admits the representation

$$\begin{cases} \frac{-1}{\overline{F}_n(t)} \int\limits_t^{\tau_F} \varphi_3(x-t)\, d\overline{F}_n(x) = \frac{1}{\widehat{\gamma}_t^{KIB}+1} \\ \frac{1}{\overline{F}_n(t)} \int\limits_t^{\tau_F} \overline{F}_n(x)\, d\varphi_2(x-t) = \widehat{\gamma}_t^{KIB} \end{cases}$$

Therefore, for convenience, we set

$$\begin{cases} \frac{-1}{\overline{F}(t)} \int\limits_t^{\tau_F} \varphi_3(x-t)\, d\overline{F}(x) = \frac{1}{\gamma+1} \\ \frac{1}{\overline{F}(t)} \int\limits_t^{\tau_F} \overline{F}(x)\, d\varphi_2(x-t) = \gamma \end{cases}$$

Until we consider the second empirical maximum likelihood equation in (12), we have

$$\frac{1}{\overline{F}_n(t)} \int_t^{\tau_F} \overline{F}_n(x)\, d\varphi_2(x-t) = \widehat{\gamma}_t^{KIB}.$$

For convenience, we set

$$\frac{1}{\overline{F}(t)} \int_t^{\tau_F} \overline{F}(x)\, d\varphi_2(x-t) = \gamma,$$

as stated in Lemma (6.1). Recall that if the convergence

$$\lim_{t\to\tau_F} \frac{\overline{F}(t+\sigma x)}{\overline{F}(t)} = (1+\gamma x)^{-1/\gamma}$$

holds for some $\sigma > 0$, with $\sigma = \gamma(t - \tau_F)$ for $\gamma < 0$, consult for instance [7], Theorem 1.2.5, p. 21. Hence, in the sequel, we shall suppose that $\sigma = \gamma(t - \tau_F)$, and

$$\widehat{\sigma}_t^{KIB} = \widehat{\gamma}_t^{KIB} (t - \widehat{\tau}_F).$$

Now, we present a procedure for estimating $\widehat{\gamma}_t^{KIB}$ and $\widehat{\sigma}_t^{KIB}$ with the aim of estimating the Weibull index through maximum likelihood. The procedure can be outlined as follows:

We identify $\widehat{\theta}_t^{KIB}$ as the unique solution to the equation $\psi_t(\theta) = 0$, where

$$\psi_t (\theta) = \left(\frac{1}{\overline{F}_n (t)} \int_t^{\tau_F} \overline{F}_n (x) \, d\varphi_2 (x - t) + 1 \right) \frac{-1}{\overline{F}_n (t)} \int_t^{\tau_F} \varphi_3 (x - t) \, d\overline{F}_n (x) - 1. \tag{13}$$

Once $\widehat{\theta}_t^{KIB}$ is obtained, we compute $\widehat{\gamma}_t^{KIB}$ as

$$\widehat{\gamma}_t^{KIB} = \frac{1}{\overline{F}_n (t)} \int_t^{\tau_F} \overline{F}_n (x) \, d\varphi_2 (x - t) \, . \tag{14}$$

For $-1 \leq \widehat{\gamma}_t^{KIB} < 0$, we then compute $(\widehat{\tau}_F)_t^{KIB}$ as

$$(\widehat{\tau}_F)_t^{KIB} = t + \frac{1}{\widehat{\theta}_t^{KIB}}, \tag{15}$$

and finally obtain $\widehat{\sigma}_t^{KIB}$ through

$$\widehat{\sigma}_t^{KIB} = \widehat{\gamma}_t^{KIB} \left(t - (\widehat{\tau}_F)_t^{KIB} \right). \tag{16}$$

Here, $\varphi_2 (x - t) = \log (1 - \theta (x - t))$ and $\varphi_3 (x - t) = (1 - \theta (x - t))^{-1}$.

Recall that in this paper we aim to establish the approximate normal behavior for all three Weibull distribution parameters. The KIB estimator in (14) corresponds to the tail value and depends on the maximum likelihood estimation of the GPD parameters under the Peak-Over-Threshold model when the threshold t is deterministic. In addition, the KIB estimator for the upper bound of the support of F is given in (15), while the estimator of the scale parameter is presented in (16).

3. Main results

In addition, we prove the approximate normal behavior of these three KIB estimators of the Weibull distribution parameters in Theorems (3.1), (3.2), and (3.3), respectively.

Theorem 3.1. *Suppose F is a distribution function with upper bound $\tau_F < \infty$. Assuming for $-1 \leq \gamma < 0$, the slowly varying function ℓ_F is normalized. Set $\sigma = \gamma(t - \tau_F)$ and let $t \to \tau_F$. Then, the KIB estimator $\widehat{\gamma}_t^{KIB}$ defined in (14) satisfies the following asymptotic normality:*

$$\sqrt{k} \left(\widehat{\gamma}_t^{KIB} - \gamma \right) \xrightarrow[k \to \infty]{} N\big(0, \gamma^2\big), \tag{17}$$

where γ is the tail index given by $\gamma = \frac{1}{\overline{F}(t)} \int_t^{\tau_F} \overline{F}(x)\, d\varphi_2(x-t)$, where $\theta = -\gamma/\sigma$, we have $\varphi_2(x-t) = \log\big(1-\theta(x-t)\big)$.

Prior to the proof of Theorem (3.1), we briefly mention the related work of Beirlant et al. [2], who studied the Peak-Over-Threshold model under random censoring. Our analysis, however, is conducted in the uncensored setting, where

$$\overline{H}(x) = \overline{F}(x), \qquad \gamma = \gamma_1, \qquad \tau_H = \tau_F.$$

This allows the asymptotic normality to be extended naturally to the uncensored case for $-1/2 < \gamma < 0$, with $\overline{H}(t) = \overline{F}(t)$, $\quad \overline{H}_n(t) = \overline{F}_n(t)$.

Proof. Based on (14), where $\frac{k}{n\,\overline{F}(t)} \xrightarrow{P} 1 \quad$ as $t \to \tau_F$, we obtain the following decomposition:

$$\begin{aligned}\sqrt{k}\left(\widehat{\gamma}_t^{KIB} - \gamma\right) &= \sqrt{\frac{n}{\overline{F}(t)}}\left(\int_t^{\tau_F} \overline{F}_n(x)\, d\varphi_2(x-t) - \int_t^{\tau_F} \overline{F}(x)\, d\varphi_2(x-t)\right) \\ &\quad -\sqrt{\frac{n}{\overline{F}(t)}}\left(\overline{F}_n(t) - \overline{F}(t)\right)\frac{1}{\overline{F}(t)}\int_t^{\tau_F} \overline{F}(x)\, d\varphi_2(x-t)\end{aligned}$$

By applying inequality (39) from Lemma (6.1), this expression reduces to

$$\begin{aligned}\sqrt{k}\left(\widehat{\gamma}_t^{KIB} - \gamma\right) &= \sqrt{\frac{n}{\overline{F}(t)}}\left(\int_t^{\tau_F} \overline{F}_n(x)\, d\varphi_2(x-t) - \int_t^{\tau_F} \overline{F}(x)\, d\varphi_2(x-t)\right) \\ &\quad -\gamma\sqrt{\frac{n}{\overline{F}(t)}}\left(\overline{F}_n(t) - \overline{F}(t)\right)\end{aligned}$$

Following the framework of Beirlant et al. [2], and in the absence of censoring, we obtain:

$$\sqrt{\frac{n}{\overline{F}(t)}}\begin{pmatrix} \left(\overline{F}_n(t) - \overline{F}(t)\right) \\ \int_t^{\tau_F} \overline{F}_n(x)\, d\varphi_2(x-t) - \int_t^{\tau_F} \overline{F}(x)\, d\varphi_2(x-t)\end{pmatrix} \to N\left(0, \begin{matrix} 1 & \gamma \\ \gamma & 2\gamma^2\end{matrix}\right)$$

By the continuous mapping theorem and linearity of the limiting distribution, we get

$$\sqrt{k}\left(\widehat{\gamma}_t^{KIB} - \gamma\right) \xrightarrow[k\to\infty]{} N\big(0, \gamma^2\big). \tag{17}$$

This completes the proof. □

Theorem 3.2.
Assuming the same conditions of Theorem (3.1), and for $-1 \le \gamma < 0$, we obtain

$$\sqrt{k}\left(\frac{\widehat{v}_t^{KIB}}{\theta} - 1\right) \to N\left(0, \theta^2\left(\gamma+1\right)^2\right), \tag{18}$$

as $k \to \infty$, where $\widehat{\theta}_t^{KIB}$ denotes the solution of $\psi_t(\theta) = 0$. Moreover, we have

$$\sqrt{k}\left(\psi_t\left(\widehat{\theta}_t^{KIB}\right)-\psi_t\left(\theta\right)\right)\to N\left(0,\gamma^2\left(\gamma+1\right)^2\right) \tag{19}$$

where ψ_t is defined in (13).

Proof.

We first establish the asymptotic behavior stated in (18) as $k\to\infty$. To this end, write

$$1-\widehat{\theta}_t^{KIB}C_j=(1-\theta C_j)\left(1-\frac{\left(\widehat{\theta}_t^{KIB}-\theta\right)C_j}{1-\theta C_j}\right)$$

By straightforward algebra, we obtain

$$\frac{1}{k}\sum_{j=1}^{n}\log\left(1-\widehat{\theta}_t^{KIB}C_j\right)1_{\{X_j>t\}}=\frac{1}{k}\sum_{j=1}^{n}\log(1-\theta C_j)\,1_{\{X_j>t\}} + \frac{1}{k}\sum_{j=1}^{n}\log\left(1-\frac{(\widehat{\theta}_t^{KIB}-\theta)C_j}{1-\theta C_j}\right)1_{\{X_j>t\}}. \tag{20}$$

From inequality (8), it follows

$$\sqrt{k}\left(\widehat{\gamma}_t^{KIB}-\gamma\right)=\frac{1}{\sqrt{k}}\sum_{j=1}^{n}\log\left(1-\frac{(\widehat{\theta}_t^{KIB}-\theta)C_j}{1-\theta C_j}\right)1_{\{X_j>t\}}. \tag{21}$$

By a first-order Taylor expansion around θ^*, we obtain the following approximation, with negligible remainder terms $O_p\left(\theta-\widehat{\theta}_t^{KIB}\right)$.

$$\theta_t^{KIB}\log\left(1-\frac{(\widehat{\theta}_t^{KIB}-\theta)C_j}{1-\theta C_j}\right)\approx\frac{(\theta-\widehat{\theta}_t^{KIB})C_j}{1-\widehat{\theta}_t^{KIB}C_j}+O_p\left(\theta-\widehat{\theta}_t^{KIB}\right). \tag{22}$$

Then,

$$\theta_t^{KIB}\sum_{j=1}^{n}\log\left(1-\frac{(\widehat{\theta}_t^{KIB}-\theta)C_j}{1-\theta C_j}\right)1_{\{X_j>t\}}=(\theta-\widehat{\theta}_t^{KIB})\sum_{j=1}^{n}\left(\frac{C_j}{1-\widehat{\theta}_t^{KIB}C_j}+O_p(1)\right)1_{\{X_j>t\}}.$$

Therefore, since $\widehat{\theta}_t^{KIB}\xrightarrow{P}\theta$ as $t\to\tau_F$, we obtain

$$\frac{1}{\sqrt{k}}\sum_{j=1}^{n}\log\left(1-\frac{(\widehat{\theta}_t^{KIB}-\theta)C_j}{1-\theta C_j}\right)1_{\{X_j>t\}}=\frac{\widehat{\theta}_t^{KIB}-\theta}{\theta^2}\cdot\frac{1}{\sqrt{k}}\sum_{j=1}^{n}\left(1-\frac{1}{1-\theta C_j}\right)1_{\{X_j>t\}} - O_p\left(\sqrt{k}(\widehat{\theta}_t^{KIB}-\theta)\right), \tag{24}$$

By substituting (24) into (21), we obtain

$$\sqrt{k}\left(\widehat{\gamma}_t^{KIB}-\gamma\right)=\frac{\sqrt{k}\left(\widehat{\theta}_t^{KIB}-\theta\right)}{\theta^2}\left(\frac{1}{k}\sum_{j=1}^{n}\left(1-\frac{1}{1-\theta C_j}\right)1_{\{X_j>t\}}-O_p(1)\right).$$

Moreover, we have

$$\frac{1}{k}\sum_{j=1}^{n}\left(1-\frac{1}{1-\theta C_j}\right)1_{\{X_j>t\}}=\frac{1}{\overline{F}_n(t)}\int_t^{\tau_F}\overline{F}_n(x)\,d\varphi_1(x-t).$$

As $t\to\tau_F$ and under Lemma (6.1), relation (21) becomes

$$\sqrt{k}\left(\widehat{\gamma}_t^{KIB}-\gamma\right)=\frac{\gamma}{\gamma+1}\cdot\frac{\sqrt{k}\,(\widehat{\theta}_t^{KIB}-\theta)}{\theta^2}. \tag{25}$$

Furthermore, as $k\to\infty$, we have

$$\sqrt{k}\left(\frac{\widehat{\theta}_t^{KIB}}{\theta}-1\right)\xrightarrow{\mathcal{D}}N\big(0,\ \theta^2(\gamma+1)^2\big)\,,$$

We now aim to prove relation (19) as $k\to\infty$, under the condition that $\psi_t\Big(\widehat{\theta}_t^{KIB}\Big)=0$, where ψ_t is defined in (13). To justify the limit in (19), we use a first-order Taylor expansion of ψ_t around θ. We have

$$\psi_t(\widehat{\theta}_t^{KIB})-\psi_t(\theta)=\psi_t'(\theta)\,(\widehat{\theta}_t^{KIB}-\theta)+o(\widehat{\theta}_t^{KIB}-\theta).$$

Since $\sqrt{k}(\widehat{\theta}_t^{KIB}-\theta)\xrightarrow{\mathcal{D}}N(0,\ \theta^2(\gamma+1)^2)$, and $\psi_t'(\theta)\to\frac{\gamma}{1+\gamma}$ as $t\to\tau_F$, it follows that

$\sqrt{k}\big(\psi_t(\widehat{\theta}_t^{KIB})-\psi_t(\theta)\big)\xrightarrow{\mathcal{D}}N(0,\ \gamma^2(1+\gamma)^2)$. We rewrite the function ψ_t in (13) as

$$\psi_t(\theta)=\left(\frac{1}{\overline{F}(t)}\int_t^{\tau_F}\overline{F}(x)\,d\varphi_2(x-t)+1\right)\frac{-1}{\overline{F}(t)}\int_t^{\tau_F}\varphi_3(x-t)\,d\overline{F}(x)-1, \tag{26}$$

By Lemma 6.1, we have $\psi_t(\theta)\xrightarrow{P}0$ as $t\to\tau_F$ for $-1\leq\gamma<0$. Under (11), we obtain

$$\begin{aligned}&\frac{1}{\overline{F}_n(t)}\int_t^{\tau_F}\overline{F}_n(x)\,d\varphi_2(x-t)-\frac{1}{\overline{F}(t)}\int_t^{\tau_F}\overline{F}(x)\,d\varphi_2(x-t)\\&=\frac{\widehat{\theta}_t^{KIB}-\theta}{\theta^2}\cdot\frac{1}{\overline{F}_n(t)}\int_t^{\tau_F}\overline{F}_n(x)\,d\varphi_1(x-t)-O_p\Big(\widehat{\theta}_t^{KIB}-\theta\Big).\end{aligned}$$

This expression may be rewritten as

$$\begin{aligned}&\left(\frac{-1}{\overline{F}_n(t)}\int_t^{\tau_F}\varphi_3(x-t)\,d\overline{F}_n(x)\right)^{-1}\Bigg[\left(\frac{1}{\overline{F}_n(t)}\int_t^{\tau_F}\overline{F}_n(x)\,d\varphi_2(x-t)+1\right)\\&\left(\frac{-1}{\overline{F}_n(t)}\int_t^{\tau_F}\varphi_3(x-t)\,d\overline{F}_n(x)\right)-1-\left(\left(\frac{1}{\overline{F}(t)}\int_t^{\tau_F}\overline{F}(x)\,d\varphi_2(x-t)+1\right)\right.\\&\left(\frac{-1}{\overline{F}_n(t)}\int_t^{\tau_F}\varphi_3(x-t)\,d\overline{F}_n(x)\right)-1\Bigg]\\&=\frac{\widehat{\theta}_t^{KIB}-\theta}{\theta^2}\cdot\frac{1}{\overline{F}_n(t)}\int_t^{\tau_F}\overline{F}_n(x)\,d\varphi_1(x-t)-O_p\Big(\widehat{\theta}_t^{KIB}-\theta\Big).\end{aligned}$$

As $\widehat{\theta}_t^{KIB} \to \theta$, and for convenience, we obtain

$$\frac{-1}{\overline{F}_n(t)} \int_t^{\tau_F} \varphi_3(x-t)\, d\overline{F}_n(x) \quad \xrightarrow{P} \quad \frac{-1}{\overline{F}(t)} \int_t^{\tau_F} \varphi_3(x-t)\, d\overline{F}(x).$$

Using (13) together with (26), we obtain

$$\begin{aligned}&\left(\frac{-1}{\overline{F}(t)} \int_t^{\tau_F} \varphi_3(x-t)\, d\overline{F}(x)\right)^{-1} \sqrt{k}\left[\psi_t(\theta) - \psi_t(\widehat{\theta}_t^{KIB})\right] \\ &= \frac{\sqrt{k}\,(\widehat{\theta}_t^{KIB} - \theta)}{\theta^2} \cdot \frac{1}{\overline{F}(t)} \int_t^{\tau_F} \overline{F}(x)\, d\varphi_1(x-t).\end{aligned}$$

and by Lemma (6.1), we obtain

$$\begin{aligned}\frac{-1}{\overline{F}(t)} \int_t^{\tau_F} \varphi_3(x-t)\, d\overline{F}(x) \quad &\xrightarrow{P} \quad \frac{1}{1+\gamma}, \\ \frac{1}{\overline{F}(t)} \int_t^{\tau_F} \overline{F}(x)\, d\varphi_1(x-t) \quad &\xrightarrow{P} \quad \frac{\gamma}{1+\gamma}.\end{aligned}$$

Finally, as $k \to \infty$, we have

$$\sqrt{k}\left[\psi_t(\theta) - \psi_t(\widehat{\theta}_t^{KIB})\right] \xrightarrow{\mathcal{D}} N\big(0,\ \gamma^2(1+\gamma)^2\big)\,.$$

□

Theorem 3.3.
In following Theorem (3.1), for $-1 \le \gamma < 0$, as $k \to \infty$, we achieve

$$\sqrt{k}\left((\widehat{\tau}_F)_t^{KIB} - \tau_F\right) \xrightarrow{\mathcal{D}} N\big(0,\ (\gamma+1)^2\big)\,,$$

where $\tau_F < \infty$ denotes the upper bound of the support of F, and $(\widehat{\tau}_F)_t^{KIB}$ is defined in (15).

Moreover, for $-1 \le \gamma < 0$ and as $k \to \infty$, we have

$$\sqrt{k}\begin{pmatrix} \widehat{\gamma}_t^{KIB} - \gamma \\ \dfrac{\widehat{\sigma}_t^{KIB}}{\sigma} - 1 \end{pmatrix} \xrightarrow{\mathcal{D}} N(0,\ \Sigma)\,,$$

with $\sigma = \gamma(t - \tau_F)$, where

$$\Sigma = \begin{pmatrix} \gamma^2 & \gamma\big(\theta(\gamma+1)+1\big) \\ \gamma\big(\theta(\gamma+1)+1\big) & \big(\theta(\gamma+1)+1\big)^2 \end{pmatrix},$$

and $\widehat{\sigma}_t^{KIB}$ is given in (16).

Proof.
Inasmuch as $(\widehat{\tau}_F)_t^{KIB}$ represents (15) and $\tau_F = t + \frac{1}{\theta}$ with $\theta = -\frac{\gamma}{\sigma}$, it follows that

$$(\widehat{\tau}_F)_t^{KIB} - \tau_F = \frac{\theta}{\left(\widehat{\theta}_t^{KIB}\right)^2}\left(\frac{\widehat{\theta}_t^{KIB}}{\theta} - \frac{\left(\widehat{\theta}_t^{KIB}\right)^2}{\theta^2}\right). \tag{27}$$

Then, since $\widehat{\theta}_t^{KIB} \xrightarrow{P} \theta$ and by (18), as $k \to \infty$ we obtain

$$\sqrt{k}\left((\widehat{\tau}_F)_t^{KIB} - \tau_F\right) \xrightarrow{\mathcal{D}} N\big(0,\ (\gamma+1)^2\big).$$

In addition, we have

$$\frac{\widehat{\sigma}_t^{KIB}}{\sigma} - 1 = \frac{1}{\gamma}\left(\widehat{\gamma}_t^{KIB} - \gamma\right) - \frac{\widehat{\gamma}_t^{KIB}}{\sigma}\left((\widehat{\tau}_F)_t^{KIB} - \tau_F\right).$$

Under (27) and (25), since $\widehat{\theta}_t^{KIB} \xrightarrow{P} \theta$ and $\widehat{\gamma}_t^{KIB} \xrightarrow{P} \gamma$, for $\sigma = \gamma(t - \tau_F)$, accordingly, we have

$$\sqrt{k}\left(\frac{\widehat{\sigma}_t^{KIB}}{\sigma} - 1\right) = \left(\frac{\theta(\gamma+1)+1}{\gamma}\right)\sqrt{k}\left(\widehat{\gamma}_t^{KIB} - \gamma\right).$$

As $k \to \infty$, and under (17), we conclude that

$$\sqrt{k}\left(\frac{\widehat{\sigma}_t^{KIB}}{\sigma} - 1\right) \xrightarrow{\mathcal{D}} N\big(0,\ (\theta(\gamma+1)+1)^2\big).$$

In particular, we obtain

$$\mathrm{Cov}\left(\sqrt{k}\left(\frac{\widehat{\sigma}_t^{KIB}}{\sigma} - 1\right), \sqrt{k}\left(\widehat{\gamma}_t^{KIB} - \gamma\right)\right) = \left(\frac{\theta(\gamma+1)+1}{\gamma}\right)\mathrm{Var}\left(\sqrt{k}\left(\widehat{\gamma}_t^{KIB} - \gamma\right)\right).$$

In the asymptotic framework $k \to \infty$, and under assumption (17), we establish that

$$\mathrm{Cov}\left(\sqrt{k}\left(\frac{\widehat{\sigma}_t^{KIB}}{\sigma} - 1\right), \sqrt{k}\left(\widehat{\gamma}_t^{KIB} - \gamma\right)\right) = \gamma\big(\theta(\gamma+1)+1\big).$$

The Cramér–Wold device is adopted to perform joint approximate normal behavior. Let $(a,b) \in \mathbb{R}^2 \setminus \{(0,0)\}$. Consider the linear combination

$$a\sqrt{k}\left(\widehat{\gamma}_t^{KIB} - \gamma\right) + b\sqrt{k}\left(\frac{\widehat{\sigma}_t^{KIB}}{\sigma} - 1\right).$$

Using the relations derived above, this combination can be stated as a simple linear function of $\sqrt{k}(\widehat{\gamma}_t^{KIB} - \gamma)$ and $\sqrt{k}(\widehat{\tau}_F^{KIB} - \tau_F)$, both of which converge in distribution to normal random variables. Consequently, the linear combination converges to a normal distribution with variance $(a,b)\,\Sigma\,(a,b)^\top$. By the Cramér–Wold device, we therefore conclude that

$$\sqrt{k}\begin{pmatrix} \widehat{\gamma}_t^{KIB} - \gamma \\ \dfrac{\widehat{\sigma}_t^{KIB}}{\sigma} - 1 \end{pmatrix} \xrightarrow{\mathcal{D}} N(0, \Sigma).$$

□

4. Simulation Studies: Applications

In what follows, we demonstrate the effectiveness of the proposed approach through both a controlled simulation study and an application to real-world data. Our focus is on estimating the Weibull tail-coefficient, a key parameter for characterizing the extrema behavior of relevant datasets.

4.1. **Simulation study based on the Weibull distribution.**

The Weibull distribution, characterized by progressively changing and non-constant and varying hazard functions, is widely used in reliability analysis and insurance modeling of material losses. Only the upper tail of the sample contains the relevant information for estimating the Weibull tail-coefficient. Several ad hoc estimation methods have been proposed, most of which rely on the inverse cumulative distribution function. As explained in Section (1) we present a new estimator for the Weibull tail-coefficient in this part that focuses on the MLE for the GPD parameters according to the peaks-over-threshold process. A logarithmic reciprocal of the extrema part of the distribution is the Weibull tail-coefficient $\alpha > 0$.

$$\overline{F}(x) = \exp(-H(x)), \quad \text{with } H(x) = -\log\big(\overline{F}(x)\big), \tag{28}$$

where $H \in RV_{1/\alpha}$ is regularly varying and satisfies (2). Accordingly, H takes the form

$$H(x) = x^{1/\alpha}\ell(x), \quad \text{as } x \to +\infty,\ \alpha > 0, \tag{29}$$

where ℓ is a progressively varied function, i.e. $\lim_{t\to+\infty} \frac{\ell(tx)}{\ell(x)} = 1$.

We address the estimation of $\alpha > 0$ defined by (28) using the MLE of the GPD parameters described in Section (1). Consider the sequence $X_{1,n}, \ldots, X_{n,n}$ denotes the sample quantiles obtained by arranging the i.i.d. variables $X_1, \ldots, X_n$ in increasing order, drawn from F.

Let k represent the count of exceedances over the threshold t, associated with the top order statistics $(X_{n-k,n}, \ldots, X_{n,n})$. In the approximate zone, $k = k_n$ is selected as a linker pattern satisfying k_n grows unbounded, whereas its proportion relative to n shrinks to zero. From [7], Lemma 3.4.1, the asymptotic tail of $(X_{n-k,n}, \ldots, X_{n,n})$ is expressed as

$$\overline{F}_t(x) = \frac{\overline{F}(x)}{\overline{F}(t)}, \quad F_t(x) = \frac{F(x) - F(t)}{1 - F(t)}, \tag{30}$$

with $\overline{F}_t(x) = 1 - F_t(x)$ for $x > t$. Substituting (30) into (28) yields

$$H(x) = -\log\big(\overline{F}_t(x)\,\overline{F}(t)\big).$$

According to Balkema and de Haan [1] (1974), $\overline{F}_t(x) \simeq \overline{F}_{\gamma,\sigma}(x)$, where $F_{\gamma,\sigma}$ is the cdf of the GPD with shape parameter γ and scale parameter σ, expressed as

$$F_{\gamma,\sigma}(x) = (1 - \theta x)^{-1/\gamma}, \quad \gamma \neq 0,\ \theta = -\frac{\gamma}{\sigma}.$$

With (29), H contains a parametric component $x^{1/\alpha}$ depending only on $\alpha > 0$ and a nonparametric part ℓ. Thus, we may take $H(x) = x^{1/\alpha}$. Then (28) becomes

$$x^{1/\alpha} = -\log\big(\overline{F}_{\gamma,\theta}(x)\,\overline{F}(t)\big).$$

Taking logarithms on both sides gives

$$\frac{1}{\alpha}\log x = \log\big(-\log\big(\overline{F}_{\gamma,\theta}(x)\,\overline{F}(t)\big)\big). \tag{33}$$

For $\overline{F}_n(t) \xrightarrow{\mathbb{P}} k/n$ as $t \to \tau_F$, and using the estimators $\widehat{\theta}_t^{KIB}$ and $\widehat{\gamma}_t^{KIB}$ defined in (7) and (8) via the Peak-Over-Threshold-MLE procedure, we define the Weibull tail-coefficient estimator as

$$\widehat{\alpha}_t^{KIB} = \frac{\frac{1}{k}\sum_{j=1}^{k}\log C_j}{\frac{1}{k}\sum_{j=1}^{k}\log\left(\frac{1}{\widehat{\gamma}_t^{KIB}}\log\left(1-\widehat{\theta}_t^{KIB}C_j\right)-\log\left(\frac{k}{n}\right)\right)}, \tag{34}$$

where $C_j = X_j - t$ for $1 \leq j \leq k$. Using Jensen's inequality and (8), we approximate

$$\frac{1}{k}\sum_{j=1}^{k}\log\left(\frac{1}{\widehat{\gamma}_t^{KIB}}\log\left(1-\widehat{\theta}_t^{KIB}C_j\right)-\log\left(\tfrac{k}{n}\right)\right) \simeq \log\left(1-\log\left(\tfrac{k}{n}\right)\right).$$

Thus,

$$\widehat{\alpha}^{KIB} = \frac{1}{T_n}\left(\frac{1}{k}\sum_{j=1}^{k}\log C_j\right), \quad T_n = \log\left(1-\log\left(\tfrac{k}{n}\right)\right), \tag{35}$$

with $1 \leq k \leq n-1$ and $C_j = X_j - t$ for $X_j > t$. If instead $C_j = X_j/t$ with $t = X_{n-k,n}$, then $\widehat{\alpha}^{KIB}$ in (35) is closely related to the Hill Pareto-type index estimator. Applying the Weibull tail coefficient estimator to small samples involves specialized estimators tailored to the upper tail, as standard MLE can be unstable. Techniques such as log-excesses improve accuracy by minimizing bias in finite samples, and $\widehat{\alpha}^{KIB}$ in (35) is based on the log-spacings between the k upper order statistics. It is approximated by the estimator of the expected value of the log-spacings, $\mathbb{E}(\log C_j)$. For very small datasets, we can use techniques such as Blom's approximation [3], which calculates the expected value of order log-spacings for $n < 21$ from a normal distribution. Ultimately, in the simulation section, we aim to determine the number k that provides the best estimator of the Weibull tail coefficient. Therefore, a simulation analysis is performed using

$$F(x) = 1 - \exp\left(-x^{1/\alpha}\right), \quad \alpha > 0, \tag{36}$$

where α is the Weibull tail coefficient. Therefore, we generate an experiment of dimension $n = 15$ following F given in (36) with parameter $\alpha = 2$, using the R Core Team software, specifically:

(1) Generate a selection at random of size $n = 15$ using the uniform distribution $p_i = \texttt{runif}(15).()$
(2) Define the complementary values $q_i = 1 - p_i$.
(3) Generate random samples of size $n = 15$ via the distribution function (36) by an inverted method, using the quantiles $x_{q_i} = (-\log q_i)^{\alpha}$.

To properly estimate $\widehat{\alpha}^{KIB}$, the Weibull tail-coefficient index defined in (35), it is crucial to determine a suitable value over k. A selection of k directly influences the accuracy of the estimator. The following table reports the estimates of $\widehat{\alpha}^{KIB}$ obtained for different levels of k, highlighting the sensitivity of the estimator to this parameter.

We can clearly see from the table that for $k = 3$, we find $\widehat{\alpha}^{KIB} = 2$ which is its true value. We also notice that if $(n/4) \leq k \leq (3n/4)$ we find that $\widehat{\alpha}^{KIB}$ given in

TABLE 1. Fifteen data points from (36) with $\alpha = 2$

0.000279	0.00913	0.0468	0.117	0.231
0.365	0.550	0.742	1.08	2.92
3.26	4.41	7.06	14.1	16.7

TABLE 2. Estimates of the Weibull tail-coefficient for $k = 1, 2, \ldots, 14$

k	1	2	3	4	5	6	7
$\widehat{\alpha}^{KIB}$	0.724	1.910	2.000	1.914	1.563	2.484	2.299
k	8	9	10	11	12	13	14
$\widehat{\alpha}^{KIB}$	2.097	2.076	2.026	2.055	1.893	1.285	-2.233

(35) takes values close to the true parameter, and this was consistently observed across several simulation runs.

4.2. **Real data application.** Numerous types of observed component and system failures have been effectively modeled using Weibull distributions, which are among the most widely known lifetime distributions. Failure statistics are typically divided into two categories: censored and complete. The real values obtained for every finding in the collection of data are known for complete data, while for censored data, some or all of the observations true values are unknown. Since we focus in this work with full data we will present an applied example based on real data that Chin et al [20] (2006), provided after conducting 50 failed tests. The data is presented in the table (3). Chin et al [20] (2006), have shown in their work that these data are unlikely to follow the Weibull distribution with two parameters. Then, by applying the MLE procedure to estimating the GPD parameter that discussed in section(2) which we mentioned at the beginning of this work. And with threshold $t = 0$, we obtain the estimator of the Weibull index is $\widehat{\gamma}^{KIB} = -0.4248901$ with the estimator of the scale parameter is $\widehat{\sigma}^{KIB} = 43.95602$ and $(\widehat{\tau}_F)_t^{KIB} = 103$. The following is the PP-plot, $\left(F_n\left((i-0.5)/n; F_{\widehat{\gamma}^{KIB},\widehat{\sigma}^{KIB}}(X_i)\right)\right.$, and the QQ-plot, $\left(F_n^{-1}\left((i-0.5)/n; X_{i,n}\right)\right.$, where $i = 1, \ldots, n$ and $n = 50$ which related to the data shown in the table(3).

PP-plot-and-QQ-plot-eps-converted-to.pdf

FIGURE 1. PP-plot and QQ-plot related to 50 data for failure

TABLE 3. Dataset of 50 failure times

0.12	0.43	0.92	1.14	1.24	1.61	1.93	2.38	4.51	5.09
6.79	7.64	8.45	11.90	11.94	13.01	13.25	14.32	17.47	18.10
18.66	19.23	24.39	25.01	26.41	26.80	27.75	29.69	29.84	31.65
32.64	35.00	40.70	42.34	43.05	43.40	44.36	45.40	48.14	49.10
49.44	51.17	58.62	60.29	72.13	72.22	72.25	72.29	85.20	89.52

5. Comments and conclusions

In this work, we focused on the extreme value index estimator for Weibull's approximation normal distribution, in the case where $-1 \leq \gamma < 0$.

In addition, we proposed a procedure based on the MLE approach for determining the endpoint τ_F and the GPD parameters via the Peaks-Over-Threshold approach when $-1 \leq \gamma < 0$. These estimators, denoted by $\widehat{\gamma}_t^{KIB}$, $(\widehat{\tau}_F)_t^{KIB}$, and $\widehat{\sigma}_t^{KIB}$, were studied under conditions (14), (15), and (16), respectively. Furthermore, we established their asymptotic normality when the threshold t is deterministic, using similar mathematical techniques to those employed by Beirlant et al. [2] (2010) in the context of censored data, and we adapted these results to the uncensored case for $\gamma \in [-1, 0[$.

We also introduced a new estimator of the Weibull tail-coefficient, defined in (35), which depends on using the MLE approach to determine the GPD values under the Peak-Over-Threshold model. This contribution provides a novel perspective on the estimation of Weibull distribution parameters. Therefore, this study represents a first step toward further research on the evaluation of Weibull distribution coefficients in reliability and failure data analysis.

Acknowledgement. The author gratefully acknowledges the anonymous reviewers for their careful reading and insightful suggestions, which have significantly contributed to improving the quality and clarity of this manuscript.

References

1. Balkema, A. A., and de Haan, L. (1974). Residual life time at great age. *The Annals of Probability*, 2(5), 792–804. https://projecteuclid.org/euclid.aop/1176996548
2. Beirlant, J., Guillou, A., and Toulemonde, G. (2010). Peaks-Over-Threshold modeling under random censoring. *Communications in Statistics Theory and Methods*, 39(7), 1158–1179. https://doi.org/10.1080/03610920902859599
3. Blom, G. (1958). Statistical Estimates and Transformed Beta Variables. Wiley, New York. https://gwern.net/doc/statistics/order/1958-blom-orderstatistics.pdf
4. Cohen, A. C. (1965). Maximum likelihood estimation in the Weibull distribution based on complete and censored samples. *Technometrics*, 7(4), 579–588. https://doi.org/10.2307/1266397

5. Dallah, D., Sulieman, H., & Alzaatreh, A. (2024). Relative range for skewed distributions: A tool for outlier detection. *Gulf Journal of Mathematics*, 16(2), 298–312. https://doi.org/10.56947/gjom.v16i2.1813
6. Davison, A. C., and Smith, R. L. (1990). Models for exceedances over high thresholds. *Journal of the Royal Statistical Society. Series B (Methodological)*, 52(3), 393–442. https://www.jstor.org/stable/2345667
7. de Haan, L., and Ferreira, A. (2006). *Extreme Value Theory: An Introduction*. Springer, New York. https://doi.org/10.1007/0-387-34471-3
8. de Haan, L., and Stadtmüller, U. (1996). Generalized regular variation of second order. *Journal of the Australian Mathematical Society (Series A)*, 61(3), 381–395. https://doi.org/10.1017/S144678870000046X
9. Drees, H., Ferreira, A., and de Haan, L. (2004). On maximum likelihood estimation of the extreme value index. *Annals of Applied Probability*, 14(3), 1179–1201. https://doi.org/10.1214/105051604000000279
10. Hassan, N. M., Rady, E. A., and Rashwan, N. I. (2021). L-moments method to estimate Gamma/Gompertz distribution based on ranked set sampling designs. *Journal of Mathematics and Computer Science*, 11(6), 8221–8239. http://dx.doi.org/10.28919/jmcs/5986
11. Kouider, M. R. (2019). Modified bisection algorithm for multiple roots of nonlinear equations with the R software. *SSRN Electronic Journal*. http://dx.doi.org/10.2139/ssrn.3451155
12. Kouider, M. R. (2019). On maximum likelihood estimates for the shape parameter of the generalized Pareto distribution. *Science Journal of Applied Mathematics and Statistics*, 7(5), 89–94. http://dx.doi.org/10.11648/j.sjams.20190705.15
13. Kouider, M. R., and Benatia, F. (2023). Modified bisection algorithm in estimating the extreme value index. *International Conference of Young Mathematicians, Institute of Mathematics NAS Ukraine*. https://www.imath.kiev.ua/~young/youngconf2023/Abstracts_2023/PS/Kouider_Benatia.pdf
14. Kouider, M. R., Idiou, N., and Benatia, F. (2023). Adaptive estimators of the generalized Pareto distribution parameters under random censorship and application. *Journal of Science and Arts*, 2, 395–412. https://www.josa.ro/docs/josa_2023_2/a_07_Mohammed_395-412_18p.pdf
15. Kouider, M. R., Idiou, N., and Benatia, F. (2023). Modified bisection algorithm in estimating the extreme value index under random censoring. *TWMS Journal of Applied and Engineering Mathematics*, 13(1), 1408–1422. https://jaem.isikun.edu.tr/web/images/articles/vol.13.no.4/13.pdf
16. Kouider, M. R., Idiou, N., and Benatia, F. (2024). Adaptive hazard estimator of the Peak-Over-Threshold model under random censoring with application. *Studies in Engineering and Exact Sciences*, 5(2), 1–24. https://doi.org/10.54021/seesv5n2-243
17. Kouider, M. R., Idiou, N., Toumi, S., and Benatia, F. (2024). Maximum lq-likelihood estimator of the heavy-tailed distribution parameter. *Croatian Review of Economic, Business and Social Statistics*, 10(2), 29–48. https://doi.org/10.62366/crebss.2024.2.003
18. Kouider, M. R., Kheireddine, S., and Benatia, F. (2024). The adaptive mean estimation of heavy-tailed distributions under random censoring. *Studies in Engineering and Exact Sciences*, 5(2), 1–31. https://doi.org/10.54021/seesv5n2-164
19. Kouider, M. R., Toumi, S., Benatia, F., and Hassan, N. M. (2025). On the maximum likelihood estimation for the transmuted extreme value distribution parameters. *Jordan Journal of Mathematics and Statistics*, 18(3), 351–367. https://jjms.yu.edu.jo/index.php/jjms/article/view/1415
20. Lai, C.-D., Murthy, D. N. P., and Xie, M. (2006). Weibull distributions and their applications. In *Springer Handbook of Engineering Statistics*. Springer, London. https://doi.org/10.1007/978-1-84628-288-1_3

21. Lo, F., Ba, D. B., & Diop, A. (2022). Maximum likelihood estimation in the generalized extreme value regression model for binary data. *Gulf Journal of Mathematics*, 12(2), 49–56. https://doi.org/10.56947/gjom.v12i2.733
22. Pickands, J. III (1975). Statistical inference using extreme order statistics. *The Annals of Statistics*, 3(1), 119–131. https://projecteuclid.org/euclid.aos/1176343003
23. Smith, R. L. (1985). Maximum likelihood estimation in a class of nonregular cases. *Biometrika*, 72(1), 67–90. https://doi.org/10.2307/2336336
24. Troshin, V. V. (2022). On the maximum domain of attraction for transformations of a normal random variable. *Journal of Mathematical Sciences*, 262(4), 537–543. https://doi.org/10.1007/s10958-022-05834-8
25. Zhou, C. (2009). Existence and consistency of the maximum likelihood estimator for the extreme value index. *Journal of Multivariate Analysis*, 100(4), 794–815. https://doi.org/10.1016/j.jmva.2008.08.009

6. Appendix

Building on the work of Beirlant et al. [2] and adapting it to complete data, we observe that the the consistency of the KIB estimator was originally established only for $-\frac{1}{2} < \gamma < 0$. Our results extend this scope to $-1 \leq \gamma < 0$, ensuring both existence and consistency of the estimator in this broader setting.

Although we did not obtain sufficient information on the approximate distribution behavior of the MLE under the Peak-Over-Threshold framework when the threshold t is deterministic and $-1 \leq \gamma \leq -\frac{1}{2}$, the method can nevertheless be applied to establish strong convergence. This, in turn, supports the asymptotic normality of the KIB estimator in the interval $-1 \leq \gamma \leq -\frac{1}{2}$.

Lemma 6.1.
Take τ_F as the right boundary of the backing of F and let $-1 \leq \gamma < 0$. If $t \to \tau_F$, then

$$\frac{1}{\overline{F}(t)} \int_t^{\tau_F} \overline{F}(x)\, d\varphi_1(x-t) \sim \frac{\gamma}{1+\gamma}, \tag{37}$$

$$\frac{-1}{\overline{F}(t)} \int_t^{\tau_F} \varphi_3(x-t)\, d\overline{F}(x) \sim \frac{1}{1+\gamma}, \tag{38}$$

and

$$\frac{1}{\overline{F}_n(t)} \int_t^{\tau_F} \overline{F}_n(x)\, d\varphi_2(x-t) \sim \gamma. \tag{39}$$

Here, the functions are defined as

$$\varphi_i(x-t) = \begin{cases} 1-(1-\theta(x-t))^{-1}, & i=1, \\ \log(1-\theta(x-t)), & i=2, \\ (1-\theta(x-t))^{-1}, & i=3. \end{cases}$$

with $\theta = -\gamma/\sigma$ and $\sigma = -\gamma(\tau_F - t)$.

Proof. We first prove the relation (37). As a result, given

$$\frac{1}{\overline{F}(t)} \int_t^{\tau_F} \overline{F}(x)\, d\varphi_1(x-t) = \int_t^{\tau_F} \frac{\overline{F}(x)}{\overline{F}(t)}\, d\left(1-\left(1+\frac{\gamma}{\sigma}(x-t)\right)^{-1}\right).$$

By using the relation $\sigma = -\gamma(\tau_F - t)$ and by applying the decomposition rule for integrals

$$\frac{1}{\overline{F}(t)} \int_t^{\tau_F} \overline{F}(x)\, d\varphi_1(x-t) = -\int_t^{\tau_F} \left(1 - \frac{\tau_F - t}{\tau_F - x}\right) \frac{d\overline{F}(x)}{\overline{F}(t)} + \frac{\overline{F}(\tau_F)}{\overline{F}(t)}$$

Recall that if $\gamma < 0$, then $\overline{F}(\tau_F) = 0$. Moreover, for $\gamma < 0$ we get

$$d\overline{F}(x) = \overline{F}(x) \left(\frac{1}{\gamma(\tau_F - x)} + \frac{\ell_F'\left(\frac{1}{(\tau_F - x)}\right)}{\ell_F\left(\frac{1}{(\tau_F - x)}\right)} \left(\frac{1}{(\tau_F - x)}\right)^2 \right) dx$$

Given that progressively changing function ℓ_F is normalized and that $x \to \tau_F$, we derive

$$\overline{F}(x)(\tau_F - x)^{-2} \ell_F'\left(\frac{1}{(\tau_F - x)}\right) \left(\ell_F\left(\frac{1}{(\tau_F - x)}\right)\right)^{-1} \to 0$$

Now, since

$$\frac{d\overline{F}(x)}{\overline{F}(t)} = \frac{1}{\gamma(\tau_F - x)} \cdot \frac{\overline{F}(x)}{\overline{F}(t)}\, dx, \tag{46}$$

we can substitute this expression into the integral obtained previously. This representation highlights the dependence on γ and the distance to the endpoint τ_F, which will be crucial in establishing the asymptotic equivalence in (37). By changing variables $\dfrac{\tau_F - t}{\tau_F - x} = y$, we obtain

$$\begin{aligned}
\frac{1}{\overline{F}(t)} \int_t^{\tau_F} \overline{F}(x)\, d\varphi_1(x-t) &= -\int_t^{\tau_F} \left(1 - \frac{\tau_F - t}{\tau_F - x}\right) \frac{d\overline{F}(x)}{\overline{F}(t)} \\
&= -\frac{1}{\gamma} \int_1^{\infty} y^{1/\gamma - 1}(1-\gamma)\, dy \\
&= \frac{\gamma}{1+\gamma}.
\end{aligned}$$

Secondly, we prove the relation (38). We define

$$\frac{1}{k} \sum_{j=1}^{n} \left(\frac{1}{1 - \theta C_j}\right) 1_{X_j > t} = \frac{-1}{\overline{F}_n(t)} \int_t^{\tau_F} \varphi_3(x-t)\, d\overline{F}_n(x)$$

Since $\overline{F}_n \xrightarrow{P} \overline{F}$, for convenience we can set

$$\frac{-1}{\overline{F}_n(t)} \int_t^{\tau_F} \varphi_3(x-t)\, d\overline{F}_n(x) \xrightarrow{P} \frac{-1}{\overline{F}(t)} \int_t^{\tau_F} \varphi_3(x-t)\, d\overline{F}(x).$$

Here, $\varphi_3(x-t) = (1 - \theta(x-t))^{-1}$ and, noting that $\varphi_3(0) = 1$, we obtain

$$\frac{-1}{\overline{F}(t)} \int_t^{\tau_F} \varphi_3(x-t)\, d\overline{F}(x) = 1 + \frac{1}{\overline{F}(t)} \int_t^{\tau_F} \overline{F}(x)\, d\varphi_3(x-t)$$

Since $d\varphi_3(x-t) = -d\varphi_1(x-t)$ and as $t \to \tau_F$, we get that

$$\frac{-1}{\overline{F}(t)} \int_t^{\tau_F} \varphi_3(x-t)\, d\overline{F}(x) \sim \frac{1}{\gamma+1}$$

Next, we apply the same technical computations to verify the relation (39). Then,

$$\begin{aligned}\frac{1}{\overline{F}(t)} \int_t^{\tau_F} \overline{F}(x)\, d\varphi_2(x-t) &= \frac{1}{\overline{F}(t)} \int_t^{\tau_F} \overline{F}(x)\, d\log\Big(1 + \frac{\gamma}{\sigma}(x-t)\Big) \\ &= -\int_1^{\infty} \frac{\overline{F}\Big(\tau_F - \frac{\tau_F - t}{y}\Big)}{\overline{F}(t)}\, y^{-1}\, dy,\end{aligned}$$

where

$$\lim_{t\to\tau_F} \frac{\overline{F}\Big(\tau_F - \frac{\tau_F - t}{y}\Big)}{\overline{F}(t)} = y^{1/\gamma} \implies \frac{1}{\overline{F}(t)} \int_t^{\tau_F} \overline{F}(x)\, d\varphi_2(x-t) = \gamma.$$

$\square$

[1] Department of Process Engineering, University of Constantine 3, Salah Boubinder, Constantine, Algeria.
Email address: nesrine.idiou@univ-constantine3.dz

[2] The Higher School of Social Security, Mohamed Saleh Mentouri, BenAknoun, Algiers, Algeria.
Email address: mr.kouider@edu.esss.dz

[3*] Department of Mathematics, University of Biskra, Mohamed Khider, Biskra, Algeria.
Email address: f.benatia@univ-biskra.dz