Revista de la Facultad de Ciencias Universidad Nacional de Colombia, Sede Medellín
V 11 N°2 julio-diciembre de 2022 • ISSN-e 2357-5749 • Research paper • Pages 39 a 53
DOI: https://doi.org/10.15446/rev.fac.cienc.v11n2.98450

FLAT LIKELIHOODS: THREE-PARAMETER WEIBULL MODEL CASE[a]

VEROSIMILITUDES PLANAS: CASO DEL MODELO WEIBULL DE TRES PARÁMETROS

JOSÉ A. MONTOYA[b*], GUDELIA FIGUEROA PRECIADO[c]

Recibido 15-09-2021, aceptado 27-01-2022, versión final 16-02-2022
Research paper

ABSTRACT: Criticisms of maximum likelihood estimation frequently occur when likelihood function shape becomes flat. Although some research have been done regarding the possible causes of a flat likelihood, more work is needed to expand our knowledge on this subject. In this paper we analyze the origin of Weibull flat likelihoods. In particular, we study the severity of the likelihood flatness by examining the limit behaviour of the relative profile likelihood for the three-parameter Weibull threshold parameter, when this parameter goes to infinity. In the cases discussed here, flat likelihoods are not only related to sample size but also to an embedded model problem. Due to the widespread use of the likelihood function in inferential statistical methods, it is important not only to identify factors that can cause flat likelihoods, but also to study the severity of this flattening, in order to develop or apply *ad hoc* statistical and computational methods for making inferences.
KEYWORDS: Flat likelihood function; threshold parameter; embedded model; GEV distribution; likelihood contours; profile likelihood function.

RESUMEN: Críticas a la estimación por máxima verosimilitud ocurren frecuentemente cuando la forma de la función de verosimilitud es plana. Aunque se ha realizado investigación respecto a las posibles causas de una verosimilitud plana, es necesario un mayor trabajo para expandir nuestro conocimiento sobre este tema. En este artículo se analiza el origen de verosimilitudes Weibull planas. En particular, se estudia la severidad de la planura de la verosimilitud a través de examinar el comportamiento del límite de la verosimilitud perfil relativa del parámetro umbral del modelo Weibull de tres parámetros, cuando este parámetro se va a infinito. En los casos que aquí se presentan, las verosimilitudes planas no están solamente relacionadas con el tamaño de la muestra sino también con un problema de modelos empotrados. Dado el amplio uso de la función de verosimilitud en métodos estadísticos inferenciales, es importante no solamente identificar los factores que pueden ocasionar verosimilitudes planas, sino también estudiar lo severo de este aplanamiento, a fin de aplicar métodos estadísticos y computacionales *ad hoc*, al realizar inferencias.
PALABRAS CLAVE: Función de verosimilitud plana; parámetro umbral; modelo empotrado; Distribución de VEG, contornos de verosimilitud; función de verosimilitud perfil.

---

[a]Montoya, J. A. & Figueroa-Preciado, G. (2022). Flat likelihoods: Three-parameter Weibull model case. *Rev. Fac. Cienc.*, 11 (**2**), 39–53. DOI: https://doi.org/10.15446/rev.fac.cienc.v11n2.98450
[b]PhD in Probability and Statistics. Statistics Professor. Department of Mathematics. University of Sonora, México.
[*]Corresponding author: arturo.montoya@unison.mx
[c]PhD in Probability and Statistics. Statistics Professor. Department of Mathematics. University of Sonora, México.

1. INTRODUCTION

Statistical methods based on likelihood have been subject to severe criticisms, generally focused on getting ridiculous estimation values (Harter & Moore, 1966; Breusch *et al.*, 1997; Berger *et al.*, 1999; Martins & Stedinger, 2000; Pewsey, 2000; Martins & Stedinger, 2001; El Adlouni *et al.*, 2007; Tumlinson, 2015). The shape of the likelihood function is usually associated with these strange results; however, abnormal likelihoods can be thought of as a symptom, not a cause; see Montoya (2008), Montoya *et al.* (2009), and Liu *et al.* (2015).

Maximum likelihood estimation criticisms occur quite often and many of them arise when the shape of the likelihood function is flat. Actually, the occurrence of flat likelihoods have been used to promote integrated likelihoods, Bayesian posteriors, penalized methods as well as new optimization methods (Berger *et al.*, 1999; Tsionas, 2001; Frery *et al.*, 2004; Li & Sudjianto, 2005; Ghosh *et al.*, 2006; Cole *et al.*, 2013; Lima & Cribari-Neto, 2019). However, the problem of flat likelihoods is complex and multifactorial. Some authors have related this problem with an overparameterization of the model, inappropriate reparameterizations, embedded models, a limited amount and quality of experimental data, and also a poor model fit when sample size is large (Cheng & Iles, 1990; Catchpole & Morgan, 1997; Raue *et al.*, 2009; Sundberg, 2010; Farcomeni, & Tardella, 2012; Kreutz *et al.*, 2013).

In this article we focus on the three-parameter Weibull model. This distribution is widely used in reliability studies to describe failure or waiting times in normal or accelerated life testing (Elmahdy & Aboutahoun, 2013; Elmahdy, 2015). It is also commonly applied in many other contexts such as in forestry, when studying tree diameters distribution (Green *et al.*, 1994); in medicine, for the study of the survival times of cancer patients (Khan *et al.*, 2011); rainfall studies in hydrology (Koutsoyiannis, 2004; Silva & Peiris, 2017; Bolívar-Cimé *et al.*, 2015), and also for the study of tides in climatology (De Haan, 1990). The three-parameter Weibull model has been also used by Deng *et al.* (2018) to study the fracture strength of brittle ceramics and other materials; for describing the voltage breakage of electric circuits (Hirose & Lai, 1997), and to analyze the time needed to complete a task, in cognitive psychology (Cousineau, *et al.*, 2002), among many other applications.

The three-parameter Weibull density function is given by

$$f(x;\alpha,\gamma,\beta) = \frac{\beta}{\gamma}\left(\frac{x-\alpha}{\gamma}\right)^{\beta-1}\exp\left[-\left(\frac{x-\alpha}{\gamma}\right)^{\beta}\right], \tag{1}$$

where $x \geq \alpha$, $\gamma \geq 0$, and $\beta \geq 0$. In this parameterization α is a threshold parameter, while γ and β are scale and shape parameters, respectively. When $\alpha = 0$, a two parameter Weibull distribution is obtained.

Many researchers have reported that maximum likelihood estimation of the threshold parameter, of a three-parameter Weibull distribution, can sometimes cause both computational and statistical difficulties. Several papers in statistical literature have devoted considerable effort to address these difficulties, usually proposing other estimation procedures, as alternatives to the maximum likelihood approach. These difficulties have been mainly linked with the *non-regular threshold problem* and the *embedded model problem*, as can be found in Smith & Naylor (1987), Cheng & Iles (1990) and Hirose & Lai (1997). The first problem occurs when working with density functions that have singularities, since these are inherited by the likelihood function; hence, for certain parameter values, likelihood becomes unbounded. When density functions are replaced by the probability of the observed sample, then the likelihood function poses no problems of this sort, as stated by Barnard (1967), Barnard & Sprott (1983), Montoya *et al.* (2009) and Liu *et al.* (2015), among others. Now, the *embedded model problem*, is characterized by a relatively flat or constant likelihood function over a wide range of values of the threshold parameter. Cheng & Iles (1990) considered that this is a model selection problem. They mentioned that a three-parameter Weibull distribution has an embedded two-parameter model (the extreme value distribution) and that its presence is the reason behind the relatively flat likelihood, which gives rise to the reported computational difficulties. On the other hand, Hirose & Lai (1997) considered that the embedded model problem occurs when having one more parameter that can not be estimated given the sample size.

In this paper we perform simulations to investigate to what extent the three-parameter Weibull model lead to flat likelihoods. We analyze the limit of the relative profile likelihood of threshold parameter α, when this parameter goes to infinity, and show that flat likelihoods occur frequently. In most cases, this limit value has an empirical distribution that assigns a negligible probability to values close to zero. Moreover, the empirical distribution of this limit value estimates that at least 12 percent of all likelihood functions that result from three-parameter Weibull model under small sample size $n = 20$ are essentially flat or constant likelihoods over a huge range of values the threshold parameter. We also show that flat likelihoods occur very frequently when data of a Gumbel model, embedded in the three-parameter Weibull distribution, are modelled by a three-parameter Weibull model. Specifically, for all sample size under study, the empirical distribution of this limit value estimates that at least 45 percent of all likelihood functions that result from three-parameter Weibull model under data of a Gumbel model are essentially flat or constant likelihood functions over a huge range of values the threshold parameter.

In Section 2 we present the three parameter Weibull likelihood and relative likelihood functions of the vector parameters, as well as the profile likelihood function of threshold parameter. Section 3 focuses on analyzing an example presented by Hirose & Lai (1997), that allows us to show that a lack of parameter identifiability based on observed data makes impossible to distinguish between parameter candidates. On the other hand, Section 4 is devoted to analyze some simulation scenarios where the frequency of occurrence of certain behaviour of the relative profile likelihood, is taken as an indicator for the degree of non-identifiability of the parameters, based on the sample. Finally, some conclusions are presented in Section 6.

2. WEIBULL LIKELIHOOD

Suppose $X_1, X_2, \cdots, X_k$ are i.i.d. random variables, distributed as a three-parameter Weibull with unknown parameters (α, γ, β). Let $\boldsymbol{x} = (x_1, x_2, \cdots, x_k)$ be realization of this sample. The likelihood function of the vector parameters (α, γ, β), based on the observed sample $\boldsymbol{x}$ and in (1), is

$$L(\alpha, \gamma, \beta) = \left(\frac{\beta}{\gamma}\right)^k \left[\prod_{i=1}^{k} \left(\frac{x_i - \alpha}{\gamma}\right)\right]^{\beta - 1} \exp\left[-\sum_{i=1}^{k} \left(\frac{x_i - \alpha}{\gamma}\right)^{\beta}\right], \tag{2}$$

for $\gamma \geq 0$, $\beta \geq 0$, and $\alpha \leq x_{(1)}$, where $x_{(1)}$ is the minimum value of the observed sample. Just for notational simplicity, sample dependence of L is not explicitly written.

The relative likelihood function of (α, γ, β) is a standardized version of (2); this takes the value of one at its maximum, the MLE of (α, γ, β),

$$R(\alpha, \gamma, \beta) = \frac{L(\alpha, \gamma, \beta)}{\max\limits_{\alpha, \gamma, \beta} L(\alpha, \gamma, \beta)} = \frac{L(\alpha, \gamma, \beta)}{L\left(\widehat{\alpha}, \widehat{\gamma}, \widehat{\beta}\right)}, \tag{3}$$

where, $\widehat{\alpha}$, $\widehat{\gamma}$, and $\widehat{\beta}$ are the MLE's of parameters α, γ, and β, respectively.

Now, assume that vector parameter (α, γ, β) can be partitioned as follows: an interest parameter ψ and a nuisance parameter λ. Then, the profile likelihood of ψ and its corresponding relative likelihood function are defined as

$$L_{\max}(\psi) = \max_{\lambda} L(\psi, \lambda), \tag{4}$$

$$R_{\max}(\psi) = \frac{L_{\max}(\psi)}{\max\limits_{\psi, \lambda} L(\psi, \lambda)} = \frac{L_{\max}(\psi)}{L_{\max}(\widehat{\psi})}. \tag{5}$$

For example, the profile likelihood and its corresponding relative likelihood function of $\psi = \alpha$ is

$$L_{\max}(\alpha) = \max_{\gamma, \beta} L(\alpha, \gamma, \beta), \tag{6}$$

$$R_{\max}(\alpha) = \frac{L_{\max}(\alpha)}{\max\limits_{\alpha, \gamma, \beta} L(\alpha, \gamma, \beta)} = \frac{L_{\max}(\alpha)}{L_{\max}(\widehat{\alpha})}. \tag{7}$$

In general, the profile or maximized likelihood is a powerful though simple inferential method for estimating a parameter of interest separately from the remaining unknown parameters of the statistical model. More details about the profile likelihood approach are described in Kalbfleisch (1985, Section 10.3), Pawitan (2001, Section 3.4), Sprott (2000, p. 66), Barndorff-Nielsen & Cox (1994, pp. 89-91), Serfling (2002, pp. 155-160), and Murphy & Van Der Vaart (2000). An important and mostly unknown aspect about the profile likelihood function is that it can be used to study and visualize various aspects of the full likelihood function, such as for instance those associated with the flatness of the likelihood surface. The following example illustrates how the graph of a profile likelihood, particularly the case of α, can reveal the flat nature of the likelihood function of (α, γ, β).

Table 1: Simulated sample from a three-parameter Weibull distribution with vector parameters $(0, 3.27, 10.16)$.

3.13	2.84	3.35	2.78	2.66	3.47	3.58	2.97	3.26	3.34
2.36	3.41	2.28	2.77	2.99	2.97	3.49	3.53	3.11	2.75

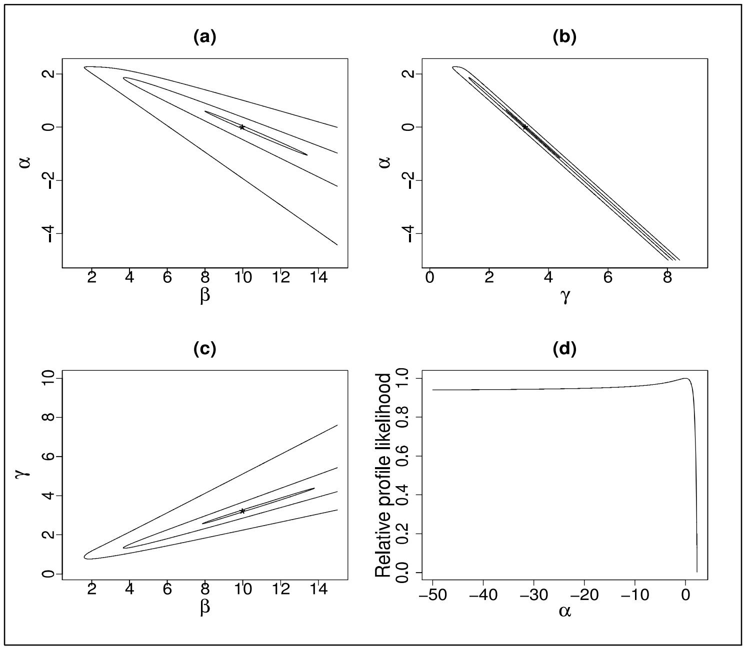

Figure 1: Relative profile likelihood functions corresponding to the data in Table 1. Source: Elaborated by the authors.

3. EXAMPLE: WEIBULL FLAT LIKELIHOOD

The dataset in Table 1 has been generated using a three-parameter Weibull distribution, with vector parameters $(\alpha, \gamma, \beta) = (0, 3.27, 10.16)$. The selection of these values was motivated by the example included in Hirose & Lai (1997), that involves $k = 20$ breakdown voltages. Figure 1(a-c) shows contour plots of the relative profile likelihood of (β, α), (γ, α), and (β, γ) for the simulated sample, displaying only three contour levels at 0.995, 0.75, and 0.05; in each one of these cases, the parameter MLE's are marked with an asterisk. It is clear from looking these elongated contours, that the three parameters are interrelated. In addition, the graph shows that the likelihood function has a ridge with a relatively flat top. In fact, Figure 1(d) shows that the relative profile likelihood of α is relatively flat over a huge range of α values.

When for a fixed data set it is observed that two different values of the parameters give rise to the same likelihood (probability of observed data), then it would be impossible to distinguish between these two parameter candidates, based on the data alone; thus, it would be not viable to identify the true parameter. In that sense, when the relative profile likelihood function of α, $L_{\max}(\alpha)$, is flat over a wide range of α values, including the MLE such as in Figure 1(d), then the probability of the observed sample is the same over all this α value set. Therefore, the lack of identifiability of these parameters occurs at least for the observed data. Since the profile likelihood function for the parameter of interest is invariant under reparameterizations of the nuisance parameters, then, the graph of the profile likelihood function of α will not change under reparameterizations of (γ,β).

Sprott (2000, p. 11) stated that the MLE and the observed information exhibit two features of the likelihood function, the MLE as a measure of its location relative to the parameter-axes, while the latter as a measure of its curvature in a neighborhood of the MLE. Thus, when the likelihood function of a scalar parameter is smooth and symmetric around the MLE, then the MLE and the observed information are usually the main features for determining the shape of the likelihood function. However, it should be stressed that flat likelihoods can not be determined by these quantities without loss of information.

Intuitively, the sensitivity of the MLE depends on the shape of the likelihood function. If the likelihood of α is very curved around MLE, this turns out to be a stable estimate of α. On the other hand, if the likelihood is flat near the MLE, this becomes highly unstable. Something important to mention, besides these facts, is that a flat likelihood can lead to likelihood-confidence intervals of α where its lower limit becomes minus infinity. In the next section we investigate the frequency of the occurrence of flat likelihoods of threshold parameter α by exploring the limit of the relative profile likelihood of α, as α approaches to minus infinity. Here we will analyze two situations: one corresponding to three-parameter Weibull samples and the other one considers samples of the embedded Gumbel model.

4. FLATNESS OF THE RELATIVE PROFILE LIKELIHOOD OF α

In this section we illustrate, to a certain degree, the non-identifiability of the parameters for the three-parameter Weibull model, considering a given sample size and focusing on the limit of the relative profile likelihood function of α, as α approaches to minus infinity,

$$R_{\max}(\infty) = \lim_{\alpha\to-\infty} R_{\max}(\alpha) = \lim_{\alpha\to-\infty} \frac{L_{\max}(\alpha)}{\max\limits_{\alpha,\gamma,\beta} L(\alpha,\gamma,\beta)}. \tag{8}$$

If $R_{\max}(\infty)$ is smaller than one, but close to it, then for every real number $\varepsilon > 0$, there is a real number α_0 such that for all $\alpha < \alpha_0$, we have $R_{\max}(\alpha) \in (R_{\max}(\infty) - \varepsilon, R_{\max}(\infty) + \varepsilon)$. Thus, there is a value set for α that makes the observed sample as likely as the MLE does. The frequency of occurrence of this type of behavior is like an indicator for the degree of non-identifiability of the parameters, based on the sample. In particular, $R_{\max}(\infty) = 1$ can be interpreted as an extreme case of lack of identifiability of the three-parameter Weibull model, based on the observed data. This situation seem to suggest that the dataset may be fitted with the embedded two-parameter model (the extreme value distribution), which corresponds to the limiting case of (1) as $\beta \to \infty$, $\gamma \to \infty$, and $\alpha \to -\infty$ such that $\gamma + \alpha \to \mu$, and $\gamma/\beta \to \sigma$ for some constants μ and σ.

Cheng & Iles (1990) showed that

$$log\left[L_{\max}(\alpha)\right] = l_0(\widehat{\mu},\widehat{\sigma}) + \frac{l_1(\widehat{\mu},\widehat{\sigma})}{\alpha} + O\left(\frac{1}{\alpha^2}\right), \tag{9}$$

where

$$l_0(\mu,\sigma) = -k\,log(\sigma) + \frac{1}{\sigma}\sum_{i=1}^{k}(x_i-\mu) - \sum_{i=1}^{k}\exp\left[\frac{(x_i-\mu)}{\sigma}\right] \tag{10}$$

is the log-likelihood function corresponding to the two-parameter extreme value model,

$$l_1(\mu,\sigma) = \sum_{i=1}^{k}(x_i-\mu) + \frac{1}{2\sigma}\sum_{i=1}^{k}(x_i-\mu)^2 - \frac{1}{2\sigma}\sum_{i=1}^{k}(x_i-\mu)^2\exp\left[\frac{(x_i-\mu)}{\sigma}\right], \tag{11}$$

and $(\widehat{\mu},\widehat{\sigma})$ is the value of (μ,σ) that maximizes $l_0(\mu,\sigma)$. Thus, for computational convenience, an approximate version of the relative profile likelihood will be used in order to calculate this limit. Our approximate version of $R_{\max}(\infty)$ is therefore

$$\widetilde{R}_{\max}(\infty) = \begin{cases} \dfrac{\exp\left[l_0(\widehat{\mu},\widehat{\sigma})\right]}{\max\limits_{\alpha,\gamma,\beta} L(\alpha,\gamma,\beta)} & \text{, if } l_1(\widehat{\mu},\widehat{\sigma}) < 0, \\ 1 & \text{, if } l_1(\widehat{\mu},\widehat{\sigma}) \geq 0. \end{cases} \tag{12}$$

4.1. Some scenarios for samples from the three-parameter Weibull distribution

An experiment was conducted to analyze the behavior of the relative profile likelihood function of α, as α approaches to minus infinity. Weibull samples of size 20, 50, 125, and 200 were simulated 10000 times, considering the following scenarios.

- Case 1, $(\alpha,\gamma,\beta) = (\alpha,4.71,21.26)$ and $\alpha = -3.46$, or -2.96 or -2.21. The selection of these values was motivated by Example 1 from Cheng & Iles (1990), concerning fibre strengths measurements, where it was found that $\widehat{\alpha} = -3.46$, $\widehat{\gamma} = 4.71$ and $\widehat{\beta} = 21.26$.
- Case 2, $(\alpha,\gamma,\beta) = (\alpha,3,10)$ and $\alpha = 0.1$, or 0.5 or 1: These parameter values were used by Hirose & Lai (1997), in order to show that the MLE of α take the value $-\infty$ with probability over 0.4, for the case of interval-grouped observations.
- Case 3, $(\alpha,\gamma,\beta) = (0.1,5.7,7),(0.1,5.7,5.5),(0.1,8,5)$. These values correspond to credible values of the posterior distribution obtained by Green *et al.* (1994) in three samples of tree diameters (samples 3-1).

Table 2 summarizes a classification of all the simulated samples for the above scenarios, according to whether $0 \leq \widetilde{R}_{\max}(\infty) < 1$ or $\widetilde{R}_{\max}(\infty) = 1$. Note that $\widetilde{R}_{\max}(\infty) = 1$ corresponds to samples where the MLE of α actually diverged to $-\infty$. At a first glance, it can be observed that as the sample size increases, there is a decrease in the proportion of samples where $\widetilde{R}_{\max}(\infty) = 1$. For all scenarios in Case 1, $\widetilde{R}_{\max}(\infty)$ assumes the value 1 with probability close to 0.36, 0.30, 0.20 and 0.15 when $n = 20$, 50, 125, and 200. For all scenarios

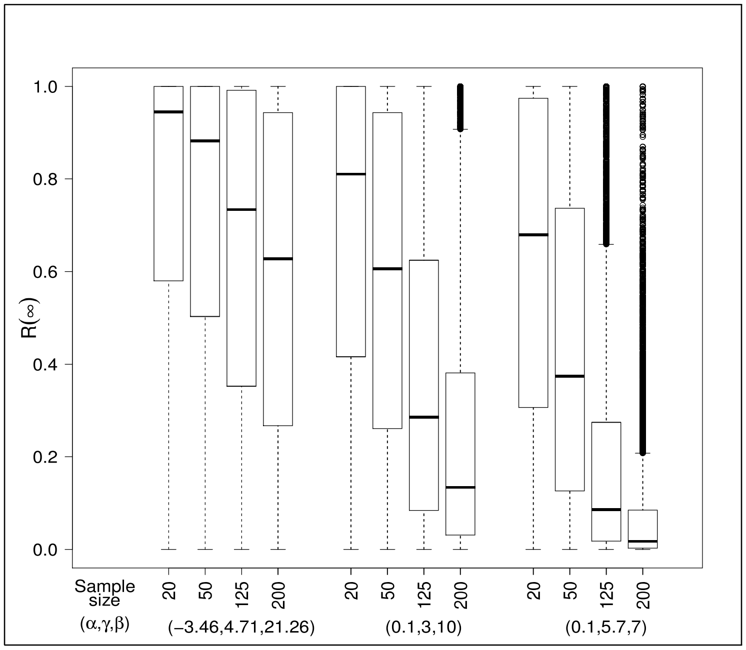

Figure 2: Box plots for three scenarios simulation Weibull of the Table 2. Source: Elaborated by the authors.

in Case 2, these probabilities are close to 0.26, 0.15, 0.04, and 0.01. In Case 3, the scenario with the highest proportion of samples where $\widetilde{R}_{\max}(\infty) = 1$ is the one where $(\alpha, \gamma, \beta) = (0.1, 5.7, 7)$, corresponding to sample 3 from Green *et al.* (1994). Here, $\widetilde{R}_{\max}(\infty)$ assumes the value 1 with probabilities close to 0.18, 0.06, 0.007 and 0, for $n = 20$, 50, 125, and 200, respectively. Thus, a general conclusion is that an embedded Gumbel model is no longer a reasonable model for the data, when sample size increases.

Box plots displayed in Figure 2 provide an idea of the shape of the frequency distribution of $\widetilde{R}_{\max}(\infty)$, for three of the simulation scenarios included in Table 2. Looking at these boxplots we can observe a large degree of non-identifiability of parameters for the three-parameter Weibull model, and notice that flat likelihoods can occur very frequently. Hence, in this kind of scenarios, it is difficult to identify the parameters of the three-parameter Weibull distribution; particularly parameter α, that represent the lower end of the support of this model.

4.2. Embedded Gumbel model in the three-parameter Weibull distribution

It is interesting to analyze scenarios such as the ones included in this section, where we simulate data from an embedded two-parameter model (the extreme value distribution) and the behavior of $\widetilde{R}_{\max}(\infty)$ is again studied. Samples of size 20, 50, 125, and 200 were simulated 10000 times from a Gumbel distribution with parameters $\mu = \alpha + \gamma$ and $\sigma = \gamma/\beta$, where the values assigned to (α, γ, β) correspond to those presented in Table 2. Once again, the samples are classified according to whether $0 \leq \widetilde{R}_{\max}(\infty) < 1$ or $\widetilde{R}_{\max}(\infty) = 1$, summarizing the results in Table 3. As we previously stated, $\widetilde{R}_{\max}(\infty) = 1$ is linked to those samples whe-

Table 2: Classification (in percentage) of 10000 simulated samples of a three-parameter Weibull random variable.

(α,γ,β)	k	$0 \leq \widetilde{R}_{\max}(\infty) < 1$	$\widetilde{R}_{\max}(\infty) = 1$
$(-3.46, 4.71, 21.26)$	20	62.24	37.76
	50	70.05	29.95
	125	79.30	20.70
	200	85.09	14.91
$(-2.96, 4.71, 21.26)$	20	63.15	36.85
	50	70.21	29.79
	125	79.06	20.94
	200	84.66	15.34
$(-2.21, 4.71, 21.26)$	20	63.98	36.02
	50	69.83	30.17
	125	79.35	20.65
	200	84.03	15.97
$(0.1, 3, 10)$	20	74.26	25.74
	50	84.66	15.34
	125	95.93	4.07
	200	98.71	1.29
$(0.5, 3, 10)$	20	73.69	26.31
	50	85.79	14.21
	125	95.66	4.34
	200	98.67	1.33
$(1, 3, 10)$	20	73.48	26.52
	50	85.67	14.33
	125	95.14	4.86
	200	98.52	1.48
$(0.1, 5.7, 7)$	20	81.59	18.41
	50	93.26	6.74
	125	99.26	0.74
	200	99.93	0.07
$(0.1, 5.7, 5.5)$	20	86.20	13.80
	50	96.92	3.08
	125	99.86	0.14
	200	100	0
$(0.1, 8, 5)$	20	88.40	11.60
	50	98.04	1.96
	125	99.92	0.08
	200	100	0

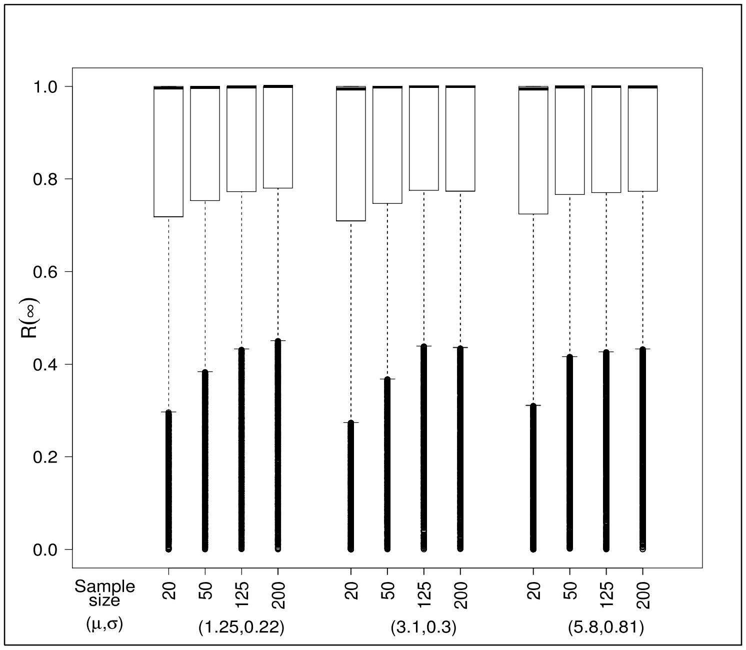

Figure 3: Box plots for three scenarios simulation Gumbel of the Table 3. Source: Elaborated by the authors.

re the MLE of α diverged to $-\infty$. These results show that in all scenarios, $\widetilde{R}_{\max}(\infty)$ assumes the value 1 with probabilities larger than 0.45. Thus, when data coming from Gumbel model, embedded in the three-parameter Weibull distribution, are modeled by a three-parameter Weibull model, flat likelihoods will occur very frequently. In order to analyze these samples in the graphical way previously presented, some box plots are displayed in Figure 3. These provide a clear idea of the shape of the frequency distribution of $\widetilde{R}_{\max}(\infty)$ for three of the simulation scenarios included in Table 3, showing that the flatness of the likelihood suggests that the embedded two-parameter model seems a reasonable model for the data.

It is important to note that once a sample size is set, for all the simulation scenarios shown in this section, the maximum computation time was two minutes, using a personal computer. On the other hand, all the simulations presented here were performed in the R software, version 4.0.3, and source code is available upon request.

Table 3: Classification (in percentage) of 10000 simulated samples of a Gumbel random variable.

(μ,σ)	k	$0 \leq \widetilde{R}_{\max}(\infty) < 1$	$\widetilde{R}_{\max}(\infty) = 1$
$(1.25, 0.22)$	20	53.40	46.60
	50	52.22	47.78
	125	51.72	48.28
	200	50	50
$(1.75, 0.22)$	20	54.38	45.62
	50	52.23	47.77
	125	51.70	48.30
	200	51.50	48.50
$(2.5, 0.22)$	20	53.02	46.98
	50	51.81	48.19
	125	51.88	48.12
	200	50.82	49.18
$(3.1, 0.3)$	20	53.83	46.17
	50	52.50	47.50
	125	51.75	48.25
	200	51.45	48.55
$(3.5, 0.3)$	20	53.81	46.19
	50	53.55	46.45
	125	51.54	48.46
	200	52.06	47.94
$(4, 0.3)$	20	53.14	46.86
	50	52.90	47.10
	125	52.22	47.78
	200	51.14	48.86
$(5.8, 0.81)$	20	53.91	46.09
	50	51.58	48.42
	125	51.85	48.15
	200	51.72	48.28
$(5.8, 1.04)$	20	53.44	46.56
	50	53.12	46.88
	125	52.23	47.77
	200	50.88	49.12
$(8.1, 1.6)$	20	53.15	46.85
	50	52.61	47.39
	125	51.71	48.29
	200	50.93	49.07

5. CONCLUSIONS

The cases discussed here were carefully selected to illustrate the potential information that flat likelihoods can provide, since their occurrence can be interpreted as a warning about an insufficient sample size for making appropriate and useful inferences about the threshold parameter, and the underlying relationship between the parameters of the model. Furthermore, flat likelihoods can also suggest that data may be coming from an embedded Gumbel model in the family of three-parameter Weibull distributions.

In general, dealing with flat likelihoods goes beyond finding stable estimators or solving numerical problems that arise during the optimization of the likelihood function, which can usually be addressed throughout an appropriate reparameterization. The flat likelihood problem embraces many issues and those should be carefully analyzed to understand the genesis of such a behaviour. The most valuable part of this paper relies on honing our intuition and understanding the multifactorial nature of this problem, contributing, in a certain way, to the scant literature concerning this subject.

In order to broaden the knowledge regarding the use of flat likelihoods in inference, future work is planned to analyze the advantages and disadvantages of using a Bayesian approach in these situations, as well as to examine the similarities and differences with the likelihood approach.

References

Barnard, G. A. (1967). The use of the likelihood function in statistical practice. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, 1, 27–40.

Barnard, G. A. & Sprott D. A. (1983). Likelihood. In: Kotz S, Johnson NL (eds) Encyclopedia of statistical science, Vol 4. Wiley, New York, pp 639–644.

Barndorff-Nielsen, O. E. & Cox, D. R. (1994). Inference and asymptotics. Chapman & Hall/CRC. Boca Raton.

Berger, J. O., Liseo, B. & Wolpert, R. L. (1999). Integrated likelihood methods for eliminating nuisance parameters. *Statistical Science*, 14(**1**), 1-28.

Bolívar-Cimé, A., Díaz-Francés, E. & Ortega, J. (2015). Optimality of profile likelihood intervals for quantiles of extreme value distributions: application to environmental disasters. *Hydrological Sciences Journal*, 60(**4**), 651-670.

Breusch, T. S., Robertson, J. C. & Welsh, A. H. (1997). The emperor's new clothes: a critique of the multivariate t regression model. *Statistica Neerlandica*, 51(**3**), 269-286.

Catchpole, E. A. & Morgan, B. J. (1997). Detecting parameter redundancy. *Biometrika*, 84(**1**), 187-196.

Cheng, R. C. H. & Iles, T. C. (1990). Embedded models in three-parameter distributions and their estimation. *Journal of the Royal Statistical Society. Series B (Methodological)*, 52(**1**), 135-149.

Cole, S.R., Chu, H. & Greenland, S. (2013). Maximum likelihood, profile likelihood, and penalized likelihood: a primer. *American Journal of Epidemiology*, 179(**2**), 252-260.

Cousineau, D., Goodman, V. W. & Shiffrin, R. M. (2002). Extending statistics of extremes to distributions varying in position and scale and the implications for race models. *Journal of Mathematical Psychology*, 46(**4**), 431-454.

De Haan, L. (1990). Fighting the arch-enemy with mathematics. *Statistica neerlandica*, 44(**2**), 45-68.

Deng, B., Jiang, D. & Gong, J. (2018). Is a three-parameter Weibull function really necessary for the characterization of the statistical variation of the strength of brittle ceramics?. *Journal of the European Ceramic Society*, 38(**4**), 2234-2242.

El Adlouni, S., Ouarda, T. B., Zhang, X., Roy, R. & Bobée, B. (2007). Generalized maximum likelihood estimators for the nonstationary generalized extreme value model. *Water Resources Research*, 43(**3**), W03410.

Elmahdy, E. E. & Aboutahoun, A. W.(2013). A new approach for parameter estimation of finite Weibull mixture distributions for reliability modeling. *Applied Mathematical Modelling*, 37(**4**), 1800-1810.

Elmahdy, E. E. (2015). A new approach for Weibull modeling for reliability life data analysis. *Applied Mathematics and computation*, 250, 708-720.

Farcomeni, A. & Tardella, L. (2012). Identifiability and inferential issues in capture-recapture experiments with heterogeneous detection probabilities. *Electronic Journal of Statistics*, 6, 2602-2626.

Frery, A. C., Cribari-Neto, F. & De Souza, M. O. (2004). Analysis of minute features in speckled imagery with maximum likelihood estimation. *EURASIP Journal on Advances in Signal Processing*, 2004(**16**), 2476-2491.

Ghosh, M., Datta, G. S., Kim, D. & Sweeting, T. J. (2006). Likelihood-based inference for the ratios of regression coefficients in linear models. *Annals of the Institute of Statistical Mathematics*, 58(**3**), 457-473.

Green, E. J., Roesch, F. A., Smith, A. F. & Strawderman, W. E. (1994). Bayesian Estimation for the Three-Parameter Weibull Distribution with Tree Diameter Data. *Biometrics*, 50(**1**), 254-269.

Harter, H. L. & Moore, A. H. (1966). Local-maximum-likelihood estimation of the parameters of three-parameter lognormal populations from complete and censored samples. *Journal of the American Statistical Association*, 61(**315**), 842-851.

Hirose, H. & Lai, T. L. (1997). Inference from grouped data in three-parameter Weibull models with applications to breakdown-voltage experiments. *Technometrics*, 39(**2**), 199-210.

Khan, H. M., Albatineh, A., Alshahrani, S., Jenkins, N. & Ahmed, N. U. (2011). Sensitivity analysis of predictive modeling for responses from the three-parameter Weibull model with a follow-up doubly censored sample of cancer patients. *Computational Statistics & Data Analysis*, 55(**12**), 3093-3103.

Kalbfleisch, J. G. (1985). Probability and Statistical Inference, Vol. 2. Springer-Verlag. New York.

Koutsoyiannis, D. (2004). Statistics of extremes and estimation of extreme rainfall: I. Theoretical investigation/Statistiques de valeurs extrêmes et estimation de précipitations extrêmes: I. Recherche théorique. *Hydrological Sciences Journal*, 49(**4**).

Kreutz, C., Raue, A., Kaschek, D. & Timmer, J. (2013). Profile likelihood in systems biology. *The FEBS journal*, 280(**11**), 2564-2571.

Li, R. & Sudjianto, A. (2005). Analysis of computer experiments using penalized likelihood in Gaussian Kriging models. *Technometrics*, 47(**2**), 111-120.

Lima, V. M. & Cribari-Neto, F. (2019). Penalized maximum likelihood estimation in the modified extended Weibull distribution. *Communications in Statistics-Simulation and Computation*, 48(**2**), 334-349.

Liu, S., Wu, H. & Meeker, W. Q. (2015). Understanding and addressing the unbounded "likelihood" problem. *The American Statistician*, 69(**3**), 191-200.

Martins, E. S. & Stedinger, J. R. (2000). Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data. *Water Resources Research*, 36(**3**), 737-744.

Martins, E. S. & Stedinger, J. R. (2001). Generalized maximum likelihood Pareto-Poisson estimators for partial duration series. *Water Resources Research*, 37(**10**), 2551-2557.

Montoya, J. A. (2008). La verosimilitud perfil en la Inferencia Estadística. Centro de Investigación en Matemáticas, A. C., Guanajuato, Gto., México.

Montoya, J. A., Díaz-Francés, E. & Sprott, D.A. (2009). On a criticism of the profile likelihood function. *Statistical Papers*, 50(**1**), 195-202.

Murphy, S. A. & Van Der Vaart, A. W. (2000). On profile likelihood. *Journal of the American Statistical Association*, 95(**450**), 449-465.

Pawitan, Y. (2001). In All Likelihood: Statistical Modelling and Inference Using Likelihood. Oxford University Press. New York.

Pewsey, A. (2000). Problems of inference for Azzalini's skewnormal distribution. *Journal of Applied Statistics*, 27(**7**), 859-870.

Raue, A., Kreutz, C., Maiwald, T., Bachmann, J., Schilling, M., Klingmüller, U. & Timmer, J. (2009). Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics*, 25(**15**), 1923-1929.

Serfling, R. J. (2002). Approximation Theorems of Mathematical Statistics. John Wiley & Sons. New York.

Silva, H. P. T. N. & Peiris, T. S. G.(2017). Statistical modeling of weekly rainfall: a case study in Colombo city in Sri Lanka. *Proceedings of the Engineering Research Conference (MERCon), Moratuwa, IEEE*, 241-246.

Smith, R. L. & Naylor, J. C. (1987). A comparison of maximum likelihood and Bayesian estimators for the three parameter Weibull distribution. *Journal of the Royal Statistical Society*, 36(**3**), 358–369.

Sprott, D. A. (2000). Statistical inference in science. Springer-Verlag. New York.

Sundberg, R. (2010). Flat and multimodal likelihoods and model lack of fit in curved exponential families. *Scandinavian Journal of Statistics*, 37(**4**), 632-643.

Tsionas, E. G. (2001). Likelihood and Posterior Shapes in Johnson's System. *Sankhyā: The Indian Journal of Statistics, Series B*, 63(**1**), 3-9.

Tumlinson, S. E. (2015). On the non-existence of maximum likelihood estimates for the extended exponential power distribution and its generalizations. *Statistics & Probability Letters*, 107, 111-114.