symmetry

MDPI

Article

Statistical Inference for the Weibull Distribution Based on δ-Record Data

Raúl Gouet [1], F. Javier López [2], Lina Maldonado [3] and Gerardo Sanz [2,*,†]

[1] Dpto. Ingeniería Matemática y Centro de Modelamiento Matemático (UMI 2807, CNRS), Universidad de Chile, Av. Libertador Bernardo O'Higgins 1058, Chile; rgouet@dim.uchile.cl

[2] Dpto. Métodos Estadísticos and BIFI, Facultad de Ciencias, Universidad de Zaragoza, 50009 Zaragoza, Spain; javier.lopez@unizar.es

[3] Departamento de Estructura e Historia Económica y Economía Pública, Facultad de Economía y Empresa, Universidad de Zaragoza, 50009 Zaragoza, Spain; lmguaje@unizar.es

* Correspondence: gerardo.sanz@unizar.es

† Current address: C/ Pedro Cerbuna 12, 50009 Zaragoza, Spain.

Received: 25 October 2019; Accepted: 16 December 2019; Published: 20 December 2019

check for updates

Abstract: We consider the maximum likelihood and Bayesian estimation of parameters and prediction of future records of the Weibull distribution from δ-record data, which consists of records and near-records. We discuss existence, consistency and numerical computation of estimators and predictors. The performance of the proposed methodology is assessed by Montecarlo simulations and the analysis of monthly rainfall series. Our conclusion is that inferences for the Weibull model, based on δ-record data, clearly improve inferences based solely on records. This methodology can be recommended, more so as near-records can be collected along with records, keeping essentially the same experimental design.

Keywords: δ-records; near-records; Weibull distribution; maximum likelihood estimation; Bayes estimation; maximum likelihood prediction; Bayes prediction

---

1. Introduction

Informally, a record is an extraordinary value of a variable, which surpasses all of its kind. Records are very popular in fields such as sports, climatology, finance or insurance. In mathematical terms, given a sequence of real-valued observations $X_1, X_2, \ldots$, we define X_1 as the first record, by convention, and we say that X_n is an upper record (or simply a record) if $X_n > M_{n-1}$ holds, where $M_{n-1} = \max\{X_1, \ldots, X_{n-1}\}$, for $n \geq 2$. Records have been extensively studied in Extreme Value Theory and their probabilistic properties, mainly under the assumption of independent and identically distributed observations, with continuous underlying distribution, are well known. This classical setting has significant symmetry which greatly simplifies calculations because it implies that all orderings of observations are equally likely. On the other hand, departures from symmetry, such as the existence of a trend in the sequence of observations, brings about technical complexities which require the use of more sophisticated mathematical tools. For general information on record theory, see Reference [1] or Reference [2].

In parallel, statistical inference for record data has developed considerably, impelled by the availability of many data sets of records and also because, in contexts such as destructive stress-testing, efficient sampling schemes (in terms of the number of broken units) yield record series. There is a vast literature on inference for record data and the interested reader can consult, for example, References [3–5].

A serious problem with record data is relative scarceness, since a sequence of n iid observations has only about $\log n$ records. So extra data may be needed and a reasonable option is near-record data, which can be available along with records. By near-record data we mean observations that are close to being records, in a sense to be made precise. Our working hypothesis is this: if one uses statistical methods specifically designed for record data but feels that the record series is too short, then a sound option is to incorporate near-records. Of course, the methodology has to be adapted to handle the new data, along with records. In this paper we show how this can be done, in the particular case of the Weibull model. We assess via simulations the impact of the additional information and present an application to real data. Various definitions of near-records have been proposed, for example, in References [6–8]. In this paper we consider near-records in the sense of Reference [7], which are closely related to δ-records. The latter were independently defined in Reference [9], as natural and tractable generalizations of records, which are as easily collected as records. Their probabilistic properties have been studied in References [9–14].

Concerning the statistical applications of δ-records, their likelihood function for a continuous distribution was first published in Reference [10], with results on maximum likelihood estimation (MLE) for the exponential and Weibull distributions. In Section 4.3 of the above cited paper, a variant of the sequential stress-testing scheme is proposed to collect δ-records, which we briefly describe here. Suppose we wish to test, say, tensile strength. In a classical sampling scheme, all items are stressed until they break. In the more efficient sequential testing, each item is stressed up to the maximum level that a previous item broke and this yields a sequence of (lower) records. The proposed variant consists in stressing the items further than the previous record, by a fixed value $\delta > 0$, to obtain a sequence of lower δ-records. The likelihood of lower δ-records can be easily obtained by adapting the ideas from the theory of standard (upper) δ-records; see Reference [10] for details. Additionally, Bayesian and MLE methods, for parameter estimation and prediction of future records in the geometric distribution, were presented in Reference [15]. A conclusion to be drawn from results in these papers is that inference methods based on δ-records outperform their corresponding record-only versions.

The main objective of the present paper is to investigate properties of inferences based on δ-records for the Weibull model, such as strong consistency of the MLE of parameters, maximum likelihood prediction (MLP) of records and Bayesian estimation and prediction of records. The reason for focusing on the Weibull distribution is twofold: First, the model and recently introduced variants are widely used in applications; see References [16,17]. Second, inference for its parameters using records has drawn significant attention in recent years; see References [4,5,18–24].

In our analysis we consider two cases: known and unknown shape parameter, while the scale parameter is always assumed unknown. According to References [25,26] and [27] (Section 14.2), in many practical problems it is not unreasonable to assume that the shape parameter is known or, at least, it is one among a small number of values. We do not analyze the situation of known scale parameter since, according to the literature, it is not considered natural and very few papers deal with it; see Section 14.2.2 in Reference [27]. It is important to warn the reader that, while statistical inference based on records from a Weibull distribution, with known shape parameter, can be reduced to inference from the exponential distribution (via power transform), this is not the case for δ-records. See Remark 2.

We assess the impact of δ-records in parameter estimation and prediction of new records, by means of Montecarlo simulation and the analysis of real data. We perform comparative analyses of estimators and predictors based on δ-records, in a variety of settings. In particular, we show that the performance of estimators and predictors is improved when using δ-record data with respect to only record data. Regarding real data, we analyze cumulative rainfall information recorded at the Castellote weather station in Spain; see Reference [10].

The paper is organized as follows: Section 2 is devoted to preliminary definitions and notation. MLE and MLP of future records are developed in Section 3; we show existence of estimators and predictors and prove the strong consistency of the MLE of the scale parameter, if the shape parameter is known. Results of simulation are presented in Section 3.4, showing that δ-records bring about noticeable improvement in estimation and prediction. The analysis of real data is presented in Section 3.5. Section 4 is devoted to Bayesian inference. We compute Bayes estimators and highest posterior density (HPD) intervals of parameters, using two different priors in the case of both parameters unknown. Then we consider Bayesian prediction of future records. Results from simulations and real data are shown in Sections 4.5 and 4.6. In Section 5 we present our conclusions.

We end this introduction with some comments about the novelty of results presented in this paper. We remark first that, while the expressions for the likelihood of the sample of δ-records and for the MLE of the parameters were first obtained in Reference [10], the remaining results are new, in the context of continuous distributions. These novel results include strong consistency of the MLE of the scale parameter (under known shape parameter); the development of a frequentist strategy for the prediction of future records and the proof of existence of predictors. In the Bayesian framework we develop the estimation of parameters, under a variety of choices of prior distributions, also complemented with a consistency result and, finally, we propose and analyze a method for predicting new records.

2. Preliminaries

Let $X_n, n \geq 1$, be a sequence of independent and identically distributed (iid) random variables, with common distribution function F_θ and density f_θ, where θ is a parameter. Let $M_n = \max\{X_1, \ldots, X_n\}$, $n \geq 1$, be the sequence of partial maxima.

Definition 1. *Let X_1 be a record by convention and, for $n \geq 2$, X_n is a (upper) record if $X_n > M_{n-1}$. The indexes $L(n)$, corresponding to record observations, are called record times. That is, $L(1) = 1$ and for $n \geq 2$, $L(n) = \min\{m > L(n-1) : X_m > M_{m-1}\}$. Records (or record-values) R_n are defined by $R_n = X_{L(n)}, n \geq 1$.*

Definition 2. *Let δ be a fixed, real parameter. Let X_1 be a δ-record by convention and, for $n \geq 2$, X_n is a δ-record if $X_n > M_{n-1} + \delta$.*

The sequences of δ-record times and δ-records are defined analogously as for records. Note that if $\delta = 0$, δ-record are just records. If $\delta > 0$, δ-records are a subsequence of records and, on the contrary, if $\delta < 0$, records are a subsequence of δ-records. So, the only statistically relevant situation is $\delta < 0$, since δ-records contain all records plus the so-called near-records (in the sense of Reference [7]). Given $\delta < 0$, X_n is a δ-near-record (or simply near-record) if $X_n \in (M_{n-1} + \delta, M_{n-1}]$. In other words, near-records are close to being records but are not records. It is clear also that near records are not symmetrically clustered around records. In the rest of the paper we assume $\delta \leq 0$.

Definition 3.
(i) A near-record X_n is said to be associated to the m-th record R_m, if $L(m) < n < L(m+1)$.
(ii) The number of near-records associated to R_m is denoted by S_m.
(iii) If $S_m > 0$, the vector of near-records associated to R_m is denoted by $(Y_{m,1}, \ldots, Y_{m,S_m})$.
(iv) The sample is defined by the vector $\mathbf{T} = (\mathbf{R}_n, \mathbf{S}_n, \mathbf{Y}_n)$, where $\mathbf{R}_n = (R_1, \ldots, R_n)$, $\mathbf{S}_n = (S_1, \ldots, S_n)$ and $\mathbf{Y}_n = (Y_{1,1}, \ldots, Y_{1,S_1}, \ldots, Y_{n,1}, \ldots, Y_{n,S_n})$.

When referring to the sample as a random object we use bold upper-case letters, otherwise we use $\mathbf{t} = (\mathbf{r}_n, \mathbf{s}_n, \mathbf{y}_n)$. Note that $\mathbf{T}$ contains a fixed number n of records (R), plus the counts (S) and respective values (Y) of all near-records associated to each record. So, $\mathbf{T}$ has random length, depending on the (random) numbers S_i of near-records associated to each record R_i. In turn, the distribution of S_i depends in a non trivial way on δ and is also affected by the tail behavior of F_θ. For example, the number of near-records obviously increases with the absolute value of δ. On the other hand, if F_θ is heavy-tailed, there will tend to be fewer near-records. Laws of large numbers and central limit theorems, for the number of near-records, can be found in References [28,29]. Normalizing sequences of asymptotic results in Example 4 of Reference [10], may serve as proxies for the expected values of the number of near-records, for the Weibull distribution.

2.1. Likelihood of δ-Record Observations

Proposition 1. *Let $\overline{F}_\theta(x) = 1 - F_\theta(x)$. The likelihood function of sample* $\mathbf{T}$ *is given by*

$$L(\mathbf{t}|\theta) = \overline{F}_\theta(r_n) \prod_{i=1}^{n} \frac{f_\theta(r_i)}{\overline{F}_\theta(r_i+\delta)^{s_i+1}} \prod_{j=1}^{s_i} f_\theta(y_{i,j}), \tag{1}$$

where $0 < r_1 < \cdots < r_n < \infty$, $s_i \in \mathbb{Z}_+ = \{0, 1, \ldots\}$ and $y_{i,j} \in (r_i + \delta, r_i)$, for $j = 1, \ldots, s_i$, $i = 1, \ldots, n$.

Proof. See Proposition 1 in Reference [10]. □

Remark 1. *Observe that in (1) the sample is assumed to contain all near-records associated to the last record R_n. To ensure that all near-records are present in the sample, the value of R_{n+1} must be observed. As commented in Section 4.2.1 of Reference [10], the data may not contain all near-records associated to R_n, since R_{n+1} is not observed. So, it is not known if some near-records associated to R_n are missing and, in such situation, the likelihood has to be modified accordingly. It is easy to see that this amounts to substituting $\overline{F}_\theta(r_n)$ for $\overline{F}_\theta(r_n+\delta)$ in (1). Then the modified likelihood $\mathcal{L}$, of a possibly incomplete sample, is*

$$\mathcal{L}(\mathbf{t}|\theta) = \overline{F}_\theta(r_n+\delta) \prod_{i=1}^{n} \frac{f_\theta(r_i)}{\overline{F}_\theta(r_i+\delta)^{s_i+1}} \prod_{j=1}^{s_i} f_\theta(y_{i,j}). \tag{2}$$

In the rest of the paper, except in the analysis of real data, we work with L (thereby assuming that all near-records associated to R_n have been collected). Results for L can be adapted to $\mathcal{L}$, with only minor modifications.

2.2. Density of Future Records

The prediction of a future record R_m, $m > n$, is based on the conditional density, given $\mathbf{T}$, as presented below.

Proposition 2. *The density of the m-th record R_m conditional on* $\mathbf{T}$, *is given by*

$$f_{R_m}(z|\mathbf{t},\theta) = \frac{(\Lambda_\theta(z) - \Lambda_\theta(r_n))^{m-n-1}}{\Gamma(m-n)} \frac{f_\theta(z)}{\overline{F}_\theta(r_n)}, \tag{3}$$

for $m > n$ and $z \geq r_n$, where $\Lambda_\theta(z) = -\log \overline{F}_\theta(z)$.

Proof. Conditionally on R_n, R_m is independent of $\mathbf{T}$ (see Proposition 1 in Reference [10]). Therefore, the density of R_m given $\mathbf{T}$ is the same as the density of R_m given R_n, which proves the result. □

2.3. Weibull Distribution

We present here the likelihoods corresponding the Weibull distribution, with two parametrizations: P_1, for classical (frequentist) inference, and P_2, for Bayesian analysis. We find it convenient to work with these two parametrizations mainly because they are found in the literature, associated respectively to classical or Bayesian inferences. See, for example, the Weibull Distribution Handbook [27], where the author argues that P_2 separates parameters and has the advantage over P_1, of simplifying the algebra in Bayesian manipulations.

Definition 4. *Let the parametrizations P_1, P_2 of the Weibull distribution be defined respectively by*

$$P_1: \quad f_{\lambda,\beta}(x) = \lambda^{-\beta}\beta x^{\beta-1} e^{-(x/\lambda)^\beta}, \quad F_{\lambda,\beta}(x) = 1 - e^{-(x/\lambda)^\beta},$$
$$P_2: \quad f_{\alpha,\beta}(x) = \alpha\beta x^{\beta-1} e^{-\alpha x^\beta}, \quad F_{\alpha,\beta}(x) = 1 - e^{-\alpha x^\beta},$$

for $x \geq 0$ and $\lambda, \alpha, \beta > 0$.

Parameter λ in P_1 is the so-called scale parameter, while β is known as shape parameter. In P_2 the scale parameter is given by $\alpha^{-1/\beta}$. From (1) we obtain the corresponding likelihoods

$$L_1(\mathbf{t}|\lambda,\beta) = \lambda^{-N\beta}\beta^N e^{-\lambda^{-\beta}G(\beta)} J^{\beta-1}, \tag{4}$$

$$L_2(\mathbf{t}|\alpha,\beta) = \alpha^N \beta^N e^{-\alpha G(\beta)} J^{\beta-1}, \tag{5}$$

where $N, G(\beta)$ and J are defined by

$$\begin{aligned} &N = n + \sum_{i=1}^{n} s_i, \quad G(\beta) = \sum_{i=1}^{n}\left(r_i^\beta - (r_i+\delta)_+^\beta + \sum_{j=1}^{s_i}\left(y_{i,j}^\beta - (r_i+\delta)_+^\beta\right)\right) + r_n^\beta \text{ and} \\ &J = \prod_{i=1}^{n} r_i \Big(\prod_{j=1}^{s_i} y_{i,j}\Big), \quad \text{with} \quad (x)_+ := \max\{x, 0\}. \end{aligned} \tag{6}$$

For possibly incomplete data, using (2), the corresponding likelihoods are (4) and (5), with $G(\beta)$ substituted by $\tilde{G}(\beta) := G(\beta) - (r_n^\beta - (r_n+\delta)_+^\beta)$.

3. Maximum Likelihood Analysis

This section is devoted to study the MLE of parameters and the MLP of future records. We begin with the MLE of parameters λ, β, related to P_1 and, as commented before, we consider two cases: β known and β unknown. Throughout this section λ is assumed unknown.

Remark 2. *Notice that if β is known, some inferences for λ can be reduced to the exponential distribution since, if X is Weibull distributed, then X^β is exponentially distributed. However, this is not so for δ-records because, for $\beta \neq 1$, the β-th power of δ-records sampled from X, are not distributed as δ-records sampled from X^β. The reason being that δ-record extraction and power transform of the data are not symmetrical actions, in the sense that they do not commute. Therefore, the consistency result of Theorem 1, for $\beta \neq 1$, does not follow from the corresponding result for the exponential model ($\beta = 1$).*

The existence of the MLE of λ, β is established in Proposition 3. For ease of notation, we omit their dependence on $\mathbf{t}$. The set of solutions of a maximization problem is denoted argmax.

3.1. MLE of Parameters λ, β

Proposition 3. *(i) For all $\beta > 0$, $\operatorname{argmax}_\lambda L_1(\mathbf{t}|\lambda,\beta)$ has a unique element*

$$\hat{\lambda}(\beta) = \left(\frac{G(\beta)}{N}\right)^{1/\beta}. \tag{7}$$

(ii) Let $\hat{A} = \operatorname{argmax}_\beta L_1\big(\mathbf{t}|\hat{\lambda}(\beta),\beta\big)$. Then $\hat{A} \neq \emptyset$ and $(\hat{\lambda}(\hat{\beta}),\hat{\beta}) \in \operatorname{argmax}_{\lambda,\beta} L_1(\mathbf{t}|\lambda,\beta)$, for any $\hat{\beta} \in \hat{A}$.

Proof. (i) (7) is obtained by solving $\frac{\partial \log L_1}{\partial \lambda} = 0$ for λ.
(ii) It can be shown that maximizing $\log L_1\big(\mathbf{t}|\hat{\lambda}(\beta),\beta\big)$ over β is equivalent to maximizing $h(\beta) := \log\beta - \log G(\beta) + (\beta-1)(\log J)/N$ and that $\emptyset \neq \operatorname{argmax}_\beta h(\beta) \subseteq [L, U]$, where $0 < L = \big(\log r_n - \frac{\log J}{N}\big)^{-1}/(3N+2)$ and $U = 2\big(\log r_n - \frac{\log J}{N}\big)^{-1}(1+\log(N+1))$. See Reference [10] for details. □

Remark 3. *Note that $\hat{\lambda}(\beta)$ depends on all δ-records (records and near-records) in the sample* $\mathbf{t}$. *If, by chance, no near-records are observed ($s_i = 0$, for $i = 1, \ldots, n$), then*

$$\hat{\lambda}(\beta) = \left(\frac{\sum_{i=1}^{n} (r_i^\beta - (r_i + \delta)_+^\beta) + r_n^\beta}{n} \right)^{1/\beta},$$

which is in contrast with the case $\delta = 0$, where the estimator depends only on the last record r_n.

The numerical computation of the MLE is straightforward. If β is known, the explicit formula for the MLE of λ is given in (7). If β is unknown, the numerical maximization of $h(\beta)$ over $[L, U]$ must be carried out to find $\hat{\beta}$, as explained in the proof of Proposition 3. Then $\hat{\lambda}(\hat{\beta})$ and $\hat{\beta}$ are the MLE of λ and β respectively.

3.2. Strong Consistency

We state below the strong consistency of $\hat{\lambda}(\beta)$, the MLE of λ when β is known. The proof of this result is split into several technical lemmas, presented in Section 6.1. Strong convergence as $n \to \infty$ is denoted $\xrightarrow{a.s.}$.

Theorem 1. $\hat{\lambda}(\beta) \xrightarrow{a.s.} \lambda$, *for all* $\lambda, \beta > 0$.

Remark 4. *If β is unknown, the question of consistency of $\hat{\lambda}$ remains open. Nevertheless, we have run Montecarlo simulations which suggest that consistency also holds in this case. In the left panel of Figure 1, values of the EMSE of $\hat{\lambda}$ are plotted versus n. It can be seen that $\hat{\lambda}$ appears to be consistent, for unknown β (in mean square sense). We also observe that convergence in the case of known β is much faster (lower curves). Moreover, in both situations δ-records yield an EMSE smaller than records. On the right panel of Figure 1 we can see a steep descent of the EMSE of $\hat{\beta}$, suggesting consistency of this estimator as well.*

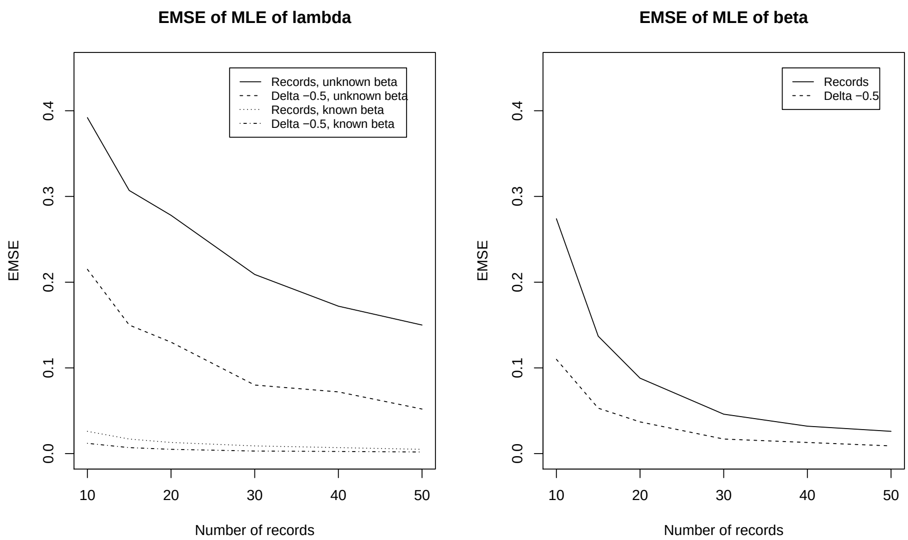

Figure 1. Simulated values of estimated mean square error (EMSE) of $\hat{\lambda}$ (left panel) and $\hat{\beta}$ (right panel), from 1000 runs, for $n \in \{10, 15, 20, 30, 40, 50\}$, $\lambda = 0.5$ and $\beta = 1$.

3.3. Maximum Likelihood Prediction of Future Records

The MLP of future records, as defined in Reference [30], consists in maximizing the so-called predictive likelihood function. The following definitions follows that idea, adapted to the sample $\mathbf{t}$ of δ-records.

Definition 5. *Let the predictive likelihood of R_m and θ be defined by $L^P(z,\mathbf{t}|\theta) = f_{R_m}(z|\mathbf{t},\theta)L(\mathbf{t}|\theta)$. In the case of the Weibull distribution, using parametrization P_1, we have, for $z \geq r_n$,*

$$f_{R_m}(z|\mathbf{t},\lambda,\beta) = \frac{\lambda^{-\beta(m-n)}\beta}{\Gamma(m-n)} z^{\beta-1} e^{-\lambda^{-\beta}(z^\beta - r_n^\beta)} (z^\beta - r_n^\beta)^{m-n-1},$$

and so,

$$L_1^P(z,\mathbf{t}|\lambda,\beta) := \frac{\lambda^{-(m-n+N)\beta}\beta^{N+1}}{\Gamma(m-n)} z^{\beta-1} e^{-\lambda^{-\beta}(z^\beta - r_n^\beta + G(\beta))} (z^\beta - r_n^\beta)^{m-n-1} J^{\beta-1}. \tag{8}$$

Definition 6. *The MLP of $R_m, m > n$, is defined by $\tilde{R}_m = \tilde{z}$, where $(\tilde{z},\tilde{\theta}) \in \operatorname*{argmax}_{z,\theta} L^P(z,\mathbf{t}|\theta)$. In the case of the Weibull distribution, using parametrization P_1, we have*
(i) $\tilde{R}_m = \tilde{z}(\beta)$, where $(\tilde{z}(\beta),\tilde{\lambda}(\beta)) \in \operatorname*{argmax}_{z,\lambda} L_1^P(z,\mathbf{t}|\lambda,\beta)$, if β is known, and
(ii) $\tilde{R}_m = \tilde{z}$, where $(\tilde{z},\tilde{\lambda},\tilde{\beta}) \in \operatorname*{argmax}_{z,\lambda,\beta} L_1^P(z,\mathbf{t}|\lambda,\beta)$, if β is unknown.

Remark 5. *The estimators $\tilde{\lambda},\tilde{\beta}$ in Definition 6 are the so-called predictive maximum likelihood estimators of λ,β, according to Reference [7]. Their properties are not investigated in this paper.*

The existence of MLP of future records in the Weibull distribution is established in Proposition 4 and the corresponding proof is presented in Section 6.2.

Proposition 4.
(i) For all $\beta > 0$, there is a unique pair $(\tilde{z}(\beta),\tilde{\lambda}(\beta)) \in \operatorname*{argmax}_{z,\lambda} L_1^P(z,\mathbf{t}|\lambda,\beta)$, given by

$$\tilde{z}(\beta) = \left(\rho + r_n^\beta\right)^{\frac{1}{\beta}} \qquad \text{and} \qquad \tilde{\lambda}(\beta) = \left(\frac{\rho + G(\beta)}{N+m-n}\right)^{\frac{1}{\beta}}, \tag{9}$$

where $\rho = \frac{-b-\sqrt{b^2-4ac}}{2a}$, with

$$a = -(\beta N + 1), \quad b = -\beta(N+1)r_n^\beta + (\beta(m-n)-1)G(\beta) \quad \text{and} \quad c = \beta(m-n-1)r_n^\beta G(\beta). \tag{10}$$

(ii) Let $\tilde{A} = \operatorname*{argmax}_{\beta} L_1^P(\tilde{z}(\beta),\mathbf{t}|\tilde{\lambda}(\beta),\beta)$. Then $\tilde{A} \neq \emptyset$ and $(\tilde{z}(\tilde{\beta}),\tilde{\lambda}(\tilde{\beta}),\tilde{\beta}) \in \operatorname*{argmax}_{z,\lambda,\beta} L_1^P(z,\mathbf{t}|\lambda,\beta)$, for any $\tilde{\beta} \in \tilde{A}$.

Remark 6. *According to Proposition 4, if β is known, there is an explicit formula for the MLP of R_m, equal to $\tilde{z}(\beta)$ in (9). On the other hand, if β is unknown, $\tilde{z}(\beta)$ and $\tilde{\lambda}(\beta)$ are plugged in L_1^P and a maximization problem in one real variable must be solved (numerically), namely $\max_{\beta} L_1^P(\tilde{z}(\beta),\mathbf{t}|\tilde{\lambda}(\beta),\beta)$. This is straightforward since, as shown in Section 6.2, there is a compact interval containing a solution $\tilde{\beta}$. Finally, it suffices to replace β by $\tilde{\beta}$ in (9), to find the MLP.*

Observe, also in Proposition 4, that, if $m = n+1$, then $b < 0$ (because $G(\beta) \leq (N+1)r_n^\beta$), $c = 0$ and, consequently, $\rho = 0$. Hence, regardless of β being known or unknown, $\tilde{R}_{n+1} = r_n$. See Section 6.2 for details.

3.4. Simulation Study

To assess the behavior of the MLE and MLP, we carry out Montecarlo simulations. For several values of λ,β and for $\delta = 0, -0.5$, we generate 10^4 samples of $n = 5$ records and their near-records.

The MLE of the Weibull parameters using δ-records were first studied in Section 4.1.1 of Reference [10]. A comparison of the accuracy of estimators, for $n = 10, \lambda = 1, \beta = 2$ and $\delta = 0, -0.5$ can be found in that paper. Throughout this paper we work in a different scenario, namely $n = 5$, with several values of λ,β and $\delta = 0, -0.5$.

The estimated mean square errors (EMSE) of the MLE of λ, β are computed as averages of the squared deviations of the MLE from the true values of the parameters. Results in Table 1 show that, for β known, the EMSE of $\hat{\lambda}(\beta)$ is much lower for $\delta < 0$ than for $\delta = 0$ (only records). For β unknown, the observed improvement is greater in $\hat{\beta}$ than in $\hat{\lambda}$. Regarding bias, it is known that the MLE of λ and β, from record data, are biased; see Reference [23]. In the case of δ-records our simulations show a small positive bias in the estimations of both parameters.

Table 1. Estimated Mean Square Errors (EMSE) of maximum likelihood estimators (MLE) $\hat{\lambda}, \hat{\beta}$.

		Known β		**Unknown β**			
		$\delta = 0$	$\delta = -0.5$	$\delta = 0$		$\delta = -0.5$	
λ	β	EMSE $\hat{\lambda}(\beta)$	EMSE $\hat{\lambda}(\beta)$	EMSE $\hat{\lambda}$	EMSE $\hat{\beta}$	EMSE $\hat{\lambda}$	EMSE $\hat{\beta}$
0.25	0.75	0.024	0.015	0.370	1.200	0.230	0.376
	1	0.013	0.004	0.118	1.814	0.033	0.174
	1.5	0.005	0.001	0.032	4.187	0.001	0.067
0.5	0.75	0.102	0.076	1.488	1.196	1.252	0.671
	1	0.051	0.027	0.463	1.871	0.298	0.631
	1.5	0.022	0.006	0.128	3.688	0.038	0.388

For MLP of future records we proceed as above in terms of the number of simulated samples, the number n of records and the values of λ, β and δ. The record to be predicted is R_7 (that is $m = n + 2$), which is the first interesting case for $m > 5$, since $\tilde{R}_6 = r_5$, as commented in Remark 6. For each run we simulate R_7 and compute the EMSE, as the average of squared deviations of $\tilde{R}_7$ from the simulated value of R_7. Results show that predictions based on records or on δ-records tend to underestimate R_7. For example, given $\lambda = 0.25$ and $\beta = 1$, the mean of R_7 is 1.75; for $\delta = 0$, the mean of $\hat{R}_7$ is 1.39 while, for $\delta = -0.5$, the mean of $\hat{R}_7$ is 1.47. The values of the EMSE are displayed in Table 2.

Table 2. Estimated Mean Square Errors (EMSE) of $\tilde{R}_7$.

		Known β		**Unknown β**	
λ	β	$\delta = 0$	$\delta = -0.5$	$\delta = 0$	$\delta = -0.5$
0.25	0.75	1.462	1.353	1.739	1.525
	1	0.223	0.193	0.260	0.213
	1.5	0.027	0.022	0.034	0.024
0.5	0.75	5.893	5.591	6.963	6.445
	1	0.884	0.793	1.072	0.917
	1.5	0.114	0.092	0.139	0.106

Although, in absolute terms, the advantage of δ-records appears to be greater in estimation than in prediction, it should be borne in mind that there exists a positive lower bound of the EMSE of any predictor of R_m, based on past information, even if parameters were known. This is because the optimal mean square predictor, based on past information, is the conditional expectation with respect to the past. Indeed, let $\mathcal{F}_n$ be the σ-algebra generated by $\{X_k;\ k \le L(n+1) - 1\}$ and R_m^* a predictor based on the available information before record R_{n+1} (that is, $\mathcal{F}_n$-measurable), then

$$E(R_m^* - R_m)^2 \ge E(E(R_m|\mathcal{F}_n) - R_m)^2 = E(E(R_m|R_n) - R_m)^2.$$

In particular, for the Weibull model, with $\beta = 1$, R_n is distributed as sums of n iid exponential random variables, with mean λ, hence

$$E(R_m^* - R_m)^2 \ge E(R_m - R_n - (m-n)\lambda)^2 = (m-n)\lambda^2. \tag{11}$$

For example, if $m = 7$, $n = 5$ and $\lambda = 0.5$, the bound in (11) is equal to 0.5. Therefore, when comparing the EMSE 0.884 ($\delta = 0$) versus 0.793 ($\delta = -0.5$), in the penultimate line of Table 2, the bound should be taken into account. A fair estimate of the gain is obtained by subtracting 0.5 from each quantity and computing $(0.384 - 0.293)/0.384$, which yields a 23.7% reduction, instead of just 10.3%, without such correction. An overall conclusion from Tables 1 and 2 is that estimators and predictors, based on δ-records, outperform those based on records only. This is coherent with the fact that there is more information in δ-records than in records and also with results in Reference [15], where this property was also observed in the case of the geometric distribution.

3.5. Real Data

We compute estimations and predictions for the rainfall data mentioned in the introduction. The data consists of cumulative rainfall (measured in millimeters), from September to November, recorded at the Castellote weather station in Spain, between 1927 and 2000. The complete sample of 74 values is well fitted by a Weibull distribution, as can be seen in Table 3 and Figure 1 in Reference [10]. We have also run the Ljung-Box test to check the existence of autocorrelation, yielding a p-value of 0.621. There are $n = 5$ records in the sample and, for $\delta = -25$, -50 and -75, the number of δ-records is 6, 9 and 18, respectively. For completeness we show the δ-record data in Table 3.

Table 3. Records and near-records (δ-records) from Castellote rainfall data in millimeters. Near records are shown to the right of their corresponding records (in boldface).

δ	δ-Record Values
0	**164.6, 184.9, 224.9, 247.1, 278.8**
-25	**164.6, 184.9**, 164.7, **224.9, 247.1, 278.8**
-50	**164.6**, 138.5, **184.9**, 164.7, **224.9**, 179.5, **247.1, 278.8**, 230.5
-75	**164.6**, 138.5, 108.0, **184.9**, 164.7, 130.1, **224.9**, 179.5, 154.1, 151.1, 157.3, **247.1, 278.8**, 204.3, 230.5, 214.0, 206.0, 209.1

Since in real data applications, unlike in simulations, we do not know the values of the parameters, it is not clear how to assess the impact of δ-records. This is also the case in predicting records, because no new record has been observed between the years 2000 and 2018. So, in order to measure the improvement of estimations and predictions due to δ-records, we compare our results with those obtained from the complete sample of 74 values, taken as benchmarks. Note that the complete sample can be seen as the set of δ-records with $\delta = -\infty$.

The MLE of λ and β were reported in Reference [10], showing that δ-records give estimates closer to the benchmark than the estimates from records only. For completeness we present those results here, in the second and third columns of Table 4.

Table 4. MLE $\hat{\lambda}, \hat{\beta}$ and maximum likelihood prediction (MLP) $\tilde{R}_7, \tilde{R}_8, \tilde{R}_9$ for Castellote rainfall data.

δ	$\hat{\lambda}$	$\hat{\beta}$	$\tilde{R}_7$	$\tilde{R}_8$	$\tilde{R}_9$
0	185	3.93	289.3	298.5	306.6
-25	188	3.48	294.1	306.8	317.7
-50	182	3.44	293.4	305.9	316.7
-75	150	2.90	292.7	305.1	316.3
$-\infty$	130	2.04	305.2	328.6	349.9

Maximum likelihood prediction was not considered in Reference [10]. The results in Table 4 show that the prediction with delta-records is closer to the estimation using the complete sample, even though the gain is not as clear as in the case of estimation. Recall, however, that the predicted values for future records using the whole sample are not the real values (since they have not been observed yet). Conclusions on the advantage of using delta-records over records must be drawn from

Montecarlo simulations in Section 3.4, rather than from this particular dataset. The reader interested in applications to real data can also see Section 4.2.2 in Reference [10], where the theory is adapted to lower records and lower δ-records. This is readily done by taking advantage of the symmetry between the definitions of upper and lower records.

4. Bayesian Analysis

We develop the estimation of parameters and prediction of future records in the Bayesian framework. We use parametrization P_2 because it is frequently found in the literature on Bayesian analysis of the Weibull model; see Section 14.2 of Reference [27]. As in Section 3, we analyze the cases of β known and unknown, while in this section α is assumed unknown. In all expressions below, α, β are positive. Let the Gamma(μ, p) density, with parameters $\mu, p > 0$, be denoted $\gamma(x|\mu, p) = \mu^p x^{p-1} e^{-\mu x}/\Gamma(p), x \geq 0$.

If β is known, we assume that α has prior $\pi(\alpha) = \gamma(\alpha|a, b)$, which is easily checked to be conjugate. In the case of β unknown, there seems to be no tractable conjugate family for α, β and, among several alternatives found in the literature, we decided to follow Kundu [31] and Soland [32]. In Kundu's approach α and β have independent gamma distributions; in Soland's approach, β is discrete, taking values in a finite set and α conditional on β is gamma distributed. Other options, not considered here, are found in References [33,34], where α, as well as β conditional on α, have gamma distributions. Definitions of Kundu's and Soland's priors are given below.

Definition 7. *(i) Let Kundu's prior π_1 be defined by $\pi_1(\alpha|\beta) = \gamma(\alpha|a, b)$ and $\pi_1(\beta) = \gamma(\beta|c, d)$, where a, b, c, d are hyperparameters and a, b do not depend on β.*
(ii) Let Soland's prior π_2 be defined by $\pi_2(\alpha|\beta = \beta_i) = \gamma(\alpha|a_i, b_i)$ and $\pi_2(\beta = \beta_i) = p_i$, where a_i, b_i are hyperparameters and β_i, p_i are positive known values, for $i = 1, \ldots, k$, with $\sum_{i=1}^k p_i = 1$.

For simplicity we write hereafter $\pi_2(\alpha|\beta_i), \pi_2(\beta_i)$, and so forth, instead of $\pi_2(\alpha|\beta = \beta_i), \pi_2(\beta = \beta_i)$ and so forth. In what follows we determine the posterior distributions to be used in inferences, namely $\pi(\alpha|\mathbf{t})$ if β is known and $\pi_j(\alpha, \beta|\mathbf{t}), j = 1, 2$, if β is unknown. Integrals with respect to α or β are understood on $(0, +\infty)$ and so, the limits of the integrals are omitted.

4.1. Posterior Distributions

Suppose first that β is known and recall that $\pi(\alpha) = \gamma(\alpha|a, b)$. From (5) we have $\pi(\alpha|\mathbf{t}) \propto L_2(\mathbf{t}|\alpha, \beta)\gamma(\alpha|a, b) \propto \alpha^{b+N-1} e^{-(a+G(\beta))\alpha}$ and so, $\pi(\alpha|\mathbf{t}) = \gamma(\alpha|a + G(\beta), b + N)$.

If β is unknown, using Kundu's prior we obtain the posteriors

$$
\begin{aligned}
\pi_1(\alpha, \beta|\mathbf{t}) &\propto L_2(\mathbf{t}|\alpha, \beta)\gamma(\alpha|a, b)\gamma(\beta|c, d) \propto \alpha^{b+N-1} \beta^{d+N-1} e^{-(a+G(\beta))\alpha} (J e^{-c})^{\beta}, \\
\pi_1(\alpha|\mathbf{t}) &\propto \alpha^{b+N-1} \int \beta^{d+N-1} e^{-(a+G(\beta))\alpha} (J e^{-c})^{\beta} d\beta, \\
\pi_1(\beta|\mathbf{t}) &\propto \Gamma(b + N) \frac{\beta^{d+N-1} (J e^{-c})^{\beta}}{(a + G(\beta))^{b+N}},
\end{aligned} \tag{12}
$$

with common normalizing constant (numerically computed), given by

$$
K_1 := \Gamma(b + N) \int \frac{\beta^{d+N-1} (J e^{-c})^{\beta}}{(a + G(\beta))^{b+N}} d\beta. \tag{13}
$$

For Soland's prior we have, from (5) and Definition 7, the posteriors

$$
\begin{aligned}
\pi_2(\alpha,\beta_i|\mathbf{t}) &\propto L_2(\mathbf{t}|\alpha,\beta_i)\pi_2(\alpha|\beta_i)\pi_2(\beta_i) = \alpha^{b_i+N-1}\beta_i^N e^{-(a_i+G(\beta_i))\alpha} J^{\beta_i-1}\frac{a_i^{b_i}}{\Gamma(b_i)}p_i,\\
\pi_2(\alpha|\mathbf{t}) &\propto \sum_{i=1}^{k}\alpha^{b_i+N-1}\beta_i^N e^{-(a_i+G(\beta_i))\alpha} J^{\beta_i-1}\frac{a_i^{b_i}}{\Gamma(b_i)}p_i,\\
\pi_2(\beta_i|\mathbf{t}) &\propto \frac{p_i\beta_i^N J^{\beta_i-1}}{(a_i+G(\beta_i))^{b_i+N}}\frac{a_i^{b_i}\Gamma(b_i+N)}{\Gamma(b_i)},
\end{aligned}
\tag{14}
$$

with common normalizing constant given by

$$
K_2 = \sum_{i=1}^{k}\frac{p_i\beta_i^N J^{\beta_i-1}}{(a_i+G(\beta_i))^{b_i+N}}\frac{a_i^{b_i}\Gamma(b_i+N)}{\Gamma(b_i)}. \tag{15}
$$

4.2. Estimation of Hyperparameters in Soland's Prior

A practical challenge with Soland's prior is the choice of the hyperparameters a_i, b_i. We propose to estimate them, using the empirical-Bayes-type method described below, inspired by Reference [4].

For $j = 1,\dots,n$ and $i = 1,\dots,k$, consider the expectations

$$
E(\overline{F}_{\alpha,\beta}(r_j)|\beta_i) = \int \overline{F}_{\alpha,\beta}(r_j)\pi_2(\alpha|\beta_i)d\alpha = \left(\frac{a_i}{a_i+r_j^{\beta_i}}\right)^{b_i}, \tag{16}
$$

$$
E(\overline{F}_{\alpha,\beta}(R_j)|\beta_i) = \int E(\overline{F}_{\alpha,\beta}(R_j)|\alpha,\beta_i)\pi_2(\alpha|\beta_i)d\alpha = 2^{-j}. \tag{17}
$$

Note that the value 2^{-j} in (17) follows from $\overline{F}_{\alpha,\beta}(R_j)$ being distributed as $e^{-\sum_{l=1}^{j}\xi_l}$, where $\xi_1,\dots,\xi_j$ are iid exponential, with parameter one, so $E(\overline{F}_{\alpha,\beta}(R_j)|\alpha,\beta) = E(e^{-\sum_{l=1}^{j}\xi_l}) = 2^{-j}$; see Reference [2]. Note also that (16) depends on r_j, the actual j-th record value of the sample, while (17) depends on the random variable R_j. Then a_i, b_i can be estimated by choosing, if possible (as described in Lemma 1 below), two suitable records $r_j, r_l, j < l$ and solving for a_i, b_i in the equations

$$
\left(\frac{a_i}{a_i+r_j^{\beta_i}}\right)^{b_i} = 2^{-j}, \qquad \left(\frac{a_i}{a_i+r_l^{\beta_i}}\right)^{b_i} = 2^{-l}. \tag{18}
$$

Lemma 1. *If there exist records $r_j, r_l, j < l$, such that $lr_j^{\beta_i} < jr_l^{\beta_i}$, then (18) has a solution.*

Proof. Let $x = 1/a_i$ and $y = b_i$, then (18) is equivalent to

$$
\left(1+xr_j^{\beta_i}\right)^y = 2^j, \qquad \left(1+xr_l^{\beta_i}\right)^y = 2^l. \tag{19}
$$

Solving for y in (19) and equating we have $g_j(x) := (1+xr_j^{\beta_i})^{1/j} = g_l(x) := (1+xr_l^{\beta_i})^{1/l}$. Furthermore, observe that $g_j(0) = g_l(0) = 1$ and that $g_j(x)/g_l(x) \to \infty$, as $x \to \infty$. Hence, if the derivatives satisfy $g_j'(0) < g_l'(0)$ or, equivalently, if $\left(r_j/r_l\right)^{\beta_i} < j/l$, there exists x^* such that $g_j(x^*) = g_l(x^*)$. Finally, x is replaced by x^* in either expression of (19) to solve for y. $\square$

4.3. Bayes Estimators of Parameters α, β

Bayes estimators are defined under quadratic loss. In the case of known β, the estimator of α is denoted $\hat{\alpha}_B(\beta)$ and follows at once from $\pi(\alpha|\mathbf{t}) = \gamma(\alpha|a + G(\beta), b + N)$. So,

$$\hat{\alpha}_B(\beta) = \frac{b+N}{a+G(\beta)}. \tag{20}$$

Credible intervals for α are readily obtained from $\pi(\alpha|\mathbf{t})$, as well. Additionally, it may be of interest the next result of consistency for $\hat{\alpha}_B(\beta)$, which follows from Theorem 1. See Reference [35] for a discussion on Bayesian consistency.

Corollary 1. *$\hat{\alpha}_B(\beta) \xrightarrow{a.s.} \alpha$, for all $a, b, \beta > 0$.*

Proof. From the definitions of P_1, P_2, we know that $\alpha = \lambda^{-\beta}$. Using this and (20), we can write $\hat{\alpha}_B(\beta)$ in terms of $\hat{\lambda}(\beta)$ (the MLE of λ in (7)) as $\hat{\alpha}_B(\beta) = \frac{b/N+1}{a/N+(\hat{\lambda}(\beta))^{\beta}}$. Last, Theorem 1 yields $\hat{\alpha}_B(\beta) \xrightarrow{a.s.} \lambda^{-\beta}$. $\square$

If β is unknown and Kundu's prior is used, we (numerically) compute the Bayes estimators of α, β, denoted respectively $\hat{\alpha}_B, \hat{\beta}_B$, by taking expectations of the posterior densities (12). We obtain

$$\hat{\alpha}_B = K_1^{-1}\Gamma(b+N+1)\int \frac{\beta^{d+N-1}(Je^{-c})^{\beta}}{(a+G(\beta))^{b+N+1}}d\beta, \qquad \hat{\beta}_B = K_1^{-1}\Gamma(b+N)\int \frac{\beta^{d+N}(Je^{-c})^{\beta}}{(a+G(\beta))^{b+N}}d\beta, \tag{21}$$

where K_1 is defined in (13).

In the case of Soland's prior, the Bayes estimators are also readily computed from (14), as

$$\hat{\alpha}_B = K_2^{-1}\sum_{i=1}^{k}\frac{p_i\beta_i^N a_i^{b_i} J^{\beta_i-1}}{(a_i+G(\beta_i))^{b_i+N+1}}\frac{\Gamma(b_i+N+1)}{\Gamma(b_i)}, \qquad \hat{\beta}_B = K_2^{-1}\sum_{i=1}^{k}\frac{p_i\beta_i^{N+1} a_i^{b_i} J^{\beta_i-1}}{(a_i+G(\beta_i))^{b_i+N}}\frac{\Gamma(b_i+N)}{\Gamma(b_i)}, \tag{22}$$

where K_2 is defined in (15). It should be noted that in simulations and in the analysis of real data, using Soland's prior, we prefer to estimate β using $\hat{\beta}_{\mathrm{MP}} \in \arg\max_{1\le i\le k} \pi_2(\beta_i|\mathbf{t})$, in order to stay within the set of possible β values.

4.4. Prediction of Future Records

The Bayesian prediction of future records is based either on $f_{R_m}(z|\beta, \mathbf{t})$, if β is known, or on $f_{R_m}(z|\mathbf{t})$, if β is unknown. In all densities below we assume that parameters are positive and $z \ge r_n$.

From (3) and using parametrization P_2, we have

$$f_{R_m}(z|\alpha,\beta,\mathbf{t}) = \frac{\alpha^{m-n}(z^{\beta}-r_n^{\beta})^{m-n-1}\beta z^{\beta-1}}{\Gamma(m-n)}e^{-\alpha(z^{\beta}-r_n^{\beta})}.$$

First, if β is known, we use the posterior $\pi(\alpha|\mathbf{t})$ to compute

$$\begin{aligned} f_{R_m}(z|\beta,\mathbf{t}) &= \int f_{R_m}(z|\alpha,\beta,\mathbf{t})\pi(\alpha|\mathbf{t})d\alpha \\ &= \frac{(z^{\beta}-r_n^{\beta})^{m-n-1}\beta z^{\beta-1}(a+G(\beta))^{b+N}}{\Gamma(m-n)\Gamma(b+N)}\int \alpha^{m-n+b+N-1}e^{-\alpha(z^{\beta}-r_n^{\beta}+a+G(\beta))}d\alpha \\ &= \frac{(z^{\beta}-r_n^{\beta})^{m-n-1}\beta z^{\beta-1}(a+G(\beta))^{b+N}\Gamma(b+N+m-n)}{(z^{\beta}-r_n^{\beta}+a+G(\beta))^{b+N+m-n}\Gamma(m-n)\Gamma(b+N)}. \end{aligned} \tag{23}$$

Then the Bayes predictor of R_m, given by $\hat{R}_m(\beta) = E(R_m|\beta,\mathbf{t}) = \int_{z_n}^{\infty} z f_{R_m}(z|\beta,\mathbf{t})dz$, is numerically computed as

$$\hat{R}_m(\beta) = \frac{\beta(a+G(\beta))^{b+N}\Gamma(b+N+m-n)}{\Gamma(m-n)\Gamma(b+N)} \int_{r_n}^{\infty} \frac{z^\beta (z^\beta - r_n^\beta)^{m-n-1}}{(z^\beta - r_n^\beta + a + G(\beta))^{b+N+m-n}} dz.$$

If β is unknown we use $\pi_j(\alpha,\beta)$ ($j=1$, for Kundu and $j=2$, for Soland) to compute

$$f_{R_m}(z|\mathbf{t}) = \int\int f_{R_m}(z|\alpha,\beta,\mathbf{t})\pi_j(\alpha,\beta|\mathbf{t})d\alpha d\beta, \tag{24}$$

and the Bayes predictor is given by $\hat{R}_m = E(R_m|\mathbf{t}) = \int_{z_n}^{\infty} z f_{R_m}(z|\mathbf{t})dz$. Bayesian prediction intervals are also readily obtained.

In the case of Kundu's prior, from (12) and (24) we have

$$f_{R_m}(z|\mathbf{t}) = K_1^{-1}\frac{\Gamma(N+m-n+b)}{\Gamma(m-n)} \int \frac{\beta^{N+d} z^{\beta-1}(z^\beta - r_n^\beta)^{m-n-1}(Je^{-c})^\beta}{(z^\beta - r_n^\beta + a + G(\beta))^{m-n+N+b}} d\beta. \tag{25}$$

In the case of Soland's prior, from (14) and (24), we get

$$f_{R_m}(z|\mathbf{t}) = \frac{K_2^{-1}}{\Gamma(m-n)} \sum_{i=1}^{k} \frac{\beta_i^{N+1} z^{\beta_i - 1}(z^{\beta_i} - r_n^{\beta_i})^{m-n-1} J^{\beta_i - 1} a_i^{b_i} p_i \Gamma(m-n+b_i+N)}{(z^{\beta_i} - r_n^{\beta_i} + a_i + G(\beta_i))^{m-n+N+b_i}\Gamma(b_i)}.$$

4.5. Simulation Study

We assess here the performance of estimators, credible intervals and predictors. To that end, we simulate samples of δ-records, with $n=5$ records. For each sample we compute estimators or predictors for records only ($\delta=0$) and for δ-records ($\delta=-0.5$).

4.5.1. Known β

In Table 5 we show results for the Bayes estimator $\hat{\alpha}_B(\beta)$, defined in (20) and 95% HPD intervals for α. For several values of a, b and β, we simulate 10^4 independent observations of α, from the Gamma(a,b) distribution. Then, for each α we simulate a random sample of 5 records and their associated near-records and compute $\hat{\alpha}_B(\beta)$. The EMSE is computed as average of squared differences $(\hat{\alpha}_B(\beta)-\alpha)^2$, over the 10^4 simulation runs.

Table 5. EMSE of Bayes estimator $\hat{\alpha}_B(\beta)$; length (LHPD) and coverage (%CHPD) of highest posterior density (HPD) interval of parameter α, with known β.

		$\delta=0$			$\delta=-0.5$		
$\pi(\alpha)$	β	**EMSE**	**LHPD**	**%CHPD**	**EMSE**	**LHPD**	**%CHPD**
$\gamma(\alpha\|4,4)$	0.75	0.128	1.262	94.9	0.098	1.140	95.1
	1	0.131	1.263	94.7	0.090	1.077	95.0
	1.5	0.120	1.250	94.9	0.066	0.919	94.7
$\gamma(\alpha\|4,6)$	0.75	0.224	1.719	94.9	0.162	1.500	95.0
	1	0.218	1.715	95.2	0.145	1.409	94.9
	1.5	0.222	1.719	94.9	0.107	1.206	94.6

We report in Table 5 the mean coverage and average length of the 10^4 HPD intervals. Regarding the coverage, as we sample α from its prior distribution, approximately 95% of the intervals (both for records and δ-records), contain the value of the parameter in the simulation. Since this happens in all the HPD we construct in this section, we do not include the coverage of the HPD intervals in the remaining tables. In all cases we observe that $\hat{\alpha}_B(\beta)$ and the HPD intervals based on δ-records compare quite favorably with their counterparts based only on records ($\delta=0$), in terms of smaller EMSE and interval length.

Additionally, we analyze the frequentist coverage for particular values of the parameter. In order to do so, we take $\beta = 1.5$ and fix a value of α in a grid from 0.2 to 2; we then simulate 200 samples of records (and δ-records), compute the corresponding HPDs for α, using a Gamma$(4, 4)$ prior and check if they contain the value of α. We repeat this for each value of α. Figure 2 (left) shows the coverage for $\alpha \in [0.2, 2]$. We observe that δ-records provide intervals with frequentist coverage closer to 95% than records. The right plot in Figure 2 shows the average length of the intervals as a function of α.

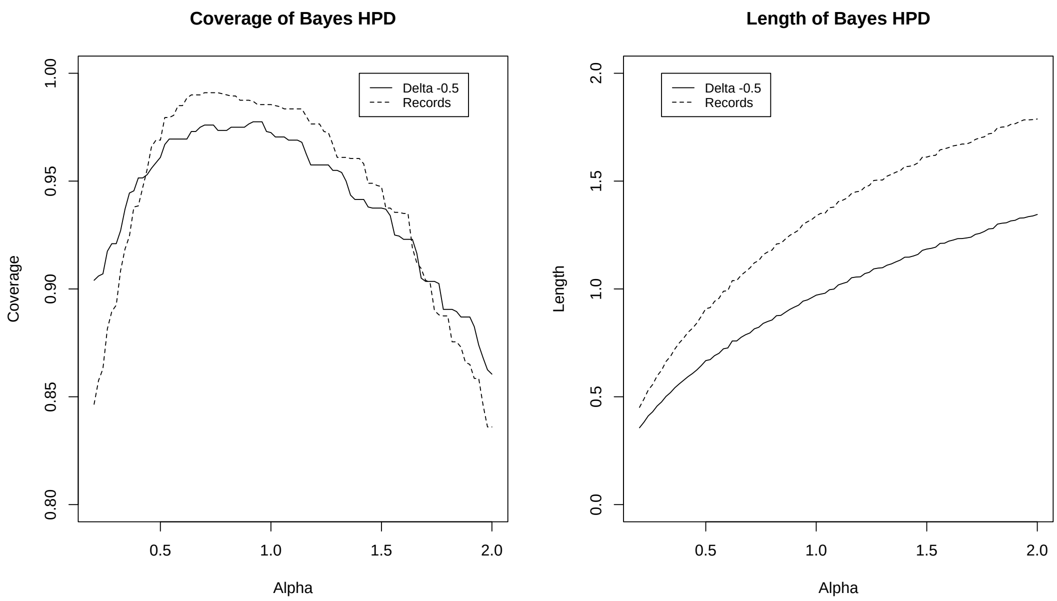

Figure 2. Frequentist coverage and average length of Bayes HPD intervals, based of records and δ-records, for several values of α, $\beta = 1.5$ and $\gamma(\alpha|4, 4)$ as α prior.

In the assessment of Bayes predictors of future records (known β), we consider different gamma priors for α and simulate 10^4 values of α. For each simulated α we generate a sample of δ-records and the values of R_6 and R_7. Then we compute the EMSE of $\hat{R}_6$ and $\hat{R}_7$, as the average of the squared deviations $(\hat{R}_6 - R_6)^2$ and $(\hat{R}_7 - R_7)^2$, over the 10^4 simulation runs. We also compute the lengths of the HPD intervals as the average of the lengths of the estimated intervals. The coverage of the intervals, defined as the proportion of runs where the simulated record is in the interval, is included as well. As in the case of estimation, since we sample α from the prior distribution in the simulations, approximately 95% of the intervals contain the corresponding record. So, we do not include the coverage in the rest of the tables for prediction.

Results are displayed in Table 6, where it is apparent that predictors are more accurate with δ-records. While Table 5 shows a significant improvement in the estimation of α, with the use of δ-records, this improvement is less visible when forecasting future records. Nevertheless, as in the case of MLP, a fair comparison between the EMSE of predictions should take into account the lower bounds commented there. For instance, if $\beta = 1$ and α is fixed, the bound is $(m-n)/\alpha^2$. Therefore, when α has a prior Gamma(a, b), with $b > 2$, the lower bound can be computed as

$$\int_0^\infty \frac{m-n}{\alpha^2} \frac{1}{\Gamma(b)} e^{-a\alpha} \alpha^{b-1} a^b d\alpha = \frac{(m-n)a^2}{(b-1)(b-2)}.$$

In the particular case $a = b = 4$, $m = 7$, $n = 5$, the bound is $16/3$. Then the gain with the use of δ-records, once subtracted the lower bound, is $(6.292 - 6.095)/(6.292 - 16/3) = 20.5\%$. That is, while the absolute gain of 0.2 may not seem relevant, relative to 6.292, it is so when the lower bound on the EMSE is taken into account.

Table 6. EMSE of Bayes predictors $\hat{R}_6, \hat{R}_7$; length (LHPD) and coverage (%CHPD) of HPD intervals for R_6, R_7, with known β.

		$\delta = 0$						$\delta = -0.5$					
		R_6			R_7			R_6			R_7		
$\pi(\alpha)$	β	EMSE $\hat{R}_6$	LHPD	%CHPD	EMSE $\hat{R}_7$	LHPD	%CHPD	EMSE $\hat{R}_6$	LHPD	%CHPD	EMSE $\hat{R}_7$	LHPD	%CHPD
	0.75	35.094	12.225	94.9	111.424	21.297	94.8	34.931	12.147	94.7	110.946	21.092	94.8
$\gamma(\alpha\|4,4)$	1	3.054	4.261	94.9	6.292	6.964	95.1	2.995	4.217	94.9	6.095	6.843	95.0
	1.5	0.227	1.368	95.4	0.491	2.140	94.6	0.219	1.335	95.3	0.454	2.058	94.9
	0.75	6.302	5.961	94.6	14.945	10.211	94.9	6.213	5.916	94.7	14.671	10.074	94.7
$\gamma(\alpha\|4,6)$	1	0.902	2.516	94.7	1.965	4.081	94.9	0.883	2.485	94.8	1.910	3.985	94.7
	1.5	0.118	0.978	95.1	0.222	1.522	94.8	0.113	0.957	94.9	0.207	1.464	94.7

4.5.2. Unknown α, β

We begin with Kundu's prior π_1, described in Definition 7. For fixed a, b, c, d we simulate 10^3 pairs (α, β) from $\pi_1(\alpha, \beta)$ and, for each (α, β) we generate a sample of 5 records and their near-records. Once the sample is observed, we numerically compute K_1 in (13), to obtain an approximation of $\pi_1(\alpha, \beta|\mathbf{t})$. We then compute the Bayes estimators $\hat{\alpha}_B, \hat{\beta}_B$ in (21) and the HPD intervals from $\pi_1(\alpha, \beta|\mathbf{t})$. The EMSE are obtained as averages of the squared deviations $(\hat{\alpha}_B - \alpha)^2$ and $(\hat{\beta}_B - \beta)^2$, over the 10^3 simulation runs. As in the case of known β, we observe in Table 7 that δ-records have a positive effect, both in the accuracy of the estimations and in the length of the HPD intervals.

Table 7. EMSE of Bayes estimators $\hat{\alpha}_B, \hat{\beta}_B$ and length (LHPD) of HPD intervals, using Kundu's prior.

		$\delta = 0$				$\delta = -0.5$			
$\pi_1(\alpha)$	$\pi_1(\beta)$	EMSE $\hat{\alpha}_B$	LHPD α	EMSE $\hat{\beta}_B$	LHPD β	EMSE $\hat{\alpha}_B$	LHPD α	EMSE $\hat{\beta}_B$	LHPD β
	$\gamma(\alpha\|8,6)$	0.179	1.488	0.048	0.731	0.152	1.389	0.036	0.662
$\gamma(\alpha\|4,4)$	$\gamma(\alpha\|6,6)$	0.192	1.485	0.080	0.970	0.160	1.363	0.058	0.850
	$\gamma(\alpha\|6,8)$	0.173	1.456	0.105	1.205	0.139	1.310	0.080	1.049
	$\gamma(\alpha\|8,6)$	0.293	1.904	0.045	0.747	0.219	1.717	0.033	0.640
$\gamma(\alpha\|4,6)$	$\gamma(\alpha\|6,6)$	0.288	1.899	0.078	0.997	0.220	1.690	0.051	0.830
	$\gamma(\alpha\|6,8)$	0.269	1.888	0.111	1.247	0.206	1.620	0.077	1.020

For Soland's prior π_2 we choose the values $\beta_k \in \{0.5, 0.75, 1, 1.25, 1.5\}$ for β, with different prior probabilities and two different gamma distributions for α, given $\beta = \beta_k$. The HPD intervals for β are not computed since β is discrete and takes only five different values. Results are shown in Table 8.

Table 8. EMSE of Bayes estimator $\hat{\alpha}_B$ and of MP estimator $\hat{\beta}_{\mathrm{MP}}$ and length (LHPD) of HPD interval for α, using Soland's prior π_2.

		$\delta = 0$			$\delta = -0.5$		
$\pi_2(\alpha\|\beta)$	$\pi_2(\beta)$	EMSE $\hat{\alpha}_B$	EMSE $\hat{\beta}_{\mathrm{MP}}$	LHPD α	EMSE $\hat{\alpha}_B$	EMSE $\hat{\beta}_{\mathrm{MP}}$	LHPD α
	$(\frac{4}{10}, \frac{2}{10}, \frac{2}{10}, \frac{1}{10}, \frac{1}{10})$	0.165	1.400	0.076	0.138	1.289	0.054
$\gamma(\alpha\|4,4)$	$(\frac{1}{10}, \frac{2}{10}, \frac{4}{10}, \frac{2}{10}, \frac{1}{10})$	0.156	1.374	0.055	0.128	1.240	0.044
	$(\frac{1}{10}, \frac{1}{10}, \frac{2}{10}, \frac{2}{10}, \frac{4}{10})$	0.163	1.354	0.079	0.121	1.166	0.065
	$(\frac{4}{10}, \frac{2}{10}, \frac{2}{10}, \frac{1}{10}, \frac{1}{10})$	0.258	1.837	0.085	0.204	1.634	0.053
$\gamma(\alpha\|4,6)$	$(\frac{1}{10}, \frac{2}{10}, \frac{4}{10}, \frac{2}{10}, \frac{1}{10})$	0.252	1.820	0.061	0.189	1.578	0.046
	$(\frac{1}{10}, \frac{1}{10}, \frac{2}{10}, \frac{2}{10}, \frac{4}{10})$	0.259	1.801	0.086	0.181	1.488	0.061

Simulation results of Bayes predictors of future records (α, β unknown), using Kundu's and Soland's priors, are presented in Tables 9 and 10, respectively. As before, we proceed by simulating first the parameter values, from the prior distributions and then the sample of δ-records and the values of future records. There is, however, a practical difficulty when computing the EMSE of predictors, since a few huge records, which actually appear in simulations, completely dominate the EMSE because it is just the average of squared deviations. Suppose, for example, that we use Kundu's prior and that the values $\alpha = 0.5, \beta = 0.1$ have been obtained from $\pi_1(\alpha, \beta)$. Then, the corresponding Weibull distribution has expectation 3.7×10^9, the value of R_m will likely be very large and so, a huge value of $(\hat{R}_m - R_m)^2$ will be observed. These "outliers" make the EMSE, computed over all the simulation runs, a useless measure of performance. In order to avoid this problem, we compute the 5% trimmed mean of $(\hat{R}_m - R_m)^2$, over the simulations. That is, we eliminate the 2.5% smallest and largest values of $(\hat{R}_m - R_m)^2$ and compute the average with the remaining ones. The results in Tables 9 and 10 show that δ-records have an impact in the prediction of future records, as observed in the case of known β.

Table 9. Trimmed EMSE (TEMSE) of Bayes predictors and lengths (LHPD) of HPD intervals for R_6, R_7, using Kundu's prior π_1.

		$\delta = 0$				$\delta = -0.5$			
$\pi_1(\alpha)$	$\pi_1(\beta)$	TEMSE $\hat{R}_6$	LHPD R_6	TEMSE $\hat{R}_7$	LHPD R_7	TEMSE $\hat{R}_6$	LHPD R_6	TEMSE $\hat{R}_7$	LHPD R_7
$\gamma(\alpha\|4,4)$	$\gamma(\alpha\|8,6)$	3318.601	42.805	24,215.125	58,844	3281.805	42.795	24198.502	58.309
	$\gamma(\alpha\|6,6)$	45.326	14.270	415.003	27.315	45.724	14.013	414.015	26.771
	$\gamma(\alpha\|6,8)$	1.997	3.735	7.296	7.446	1.917	3.602	6.885	7.288
$\gamma(\alpha\|4,6)$	$\gamma(\alpha\|8,6)$	134.211	19.175	1018.288	34.819	132.540	18.758	1006.229	34.279
	$\gamma(\alpha\|6,6)$	10.079	7.689	67.697	16.994	9.830	7.636	64.405	16.586
	$\gamma(\alpha\|6,8)$	0.664	2.266	2.539	4.221	0.702	2.201	2.534	4.032

Table 10. Trimmed EMSE (TEMSE) of Bayes predictors and lengths (LHPD) of HPD intervals for R_6, R_7, using Soland's prior π_2.

		$\delta = 0$				$\delta = -0.5$			
$\pi_2(\alpha\|\beta)$	$\pi_2(\beta)$	TEMSE $\hat{R}_6$	LHPD R_6	TEMSE $\hat{R}_7$	LHPD R_7	TEMSE $\hat{R}_6$	LHPD R_6	TEMSE $\hat{R}_7$	LHPD R_7
$\gamma(\alpha\|4,4)$	$(\frac{4}{10}, \frac{2}{10}, \frac{2}{10}, \frac{1}{10}, \frac{1}{10})$	462.880	25.479	1069.877	41.557	462.968	25.420	1054.563	41.486
	$(\frac{1}{10}, \frac{2}{10}, \frac{4}{10}, \frac{2}{10}, \frac{1}{10})$	45.904	7.823	102.881	12.856	45.608	7.714	103.401	12.634
	$(\frac{1}{10}, \frac{1}{10}, \frac{2}{10}, \frac{2}{10}, \frac{4}{10})$	15.881	5.597	56.159	10.723	16.009	5.565	54.574	10.551
$\gamma(\alpha\|4,6)$	$(\frac{4}{10}, \frac{2}{10}, \frac{2}{10}, \frac{1}{10}, \frac{1}{10})$	66.015	11.170	145.522	20.895	66.318	11.111	145.077	20.724
	$(\frac{1}{10}, \frac{2}{10}, \frac{4}{10}, \frac{2}{10}, \frac{1}{10})$	7.938	3.929	9.609	7.093	7.832	3.888	9.520	7.013
	$(\frac{1}{10}, \frac{1}{10}, \frac{2}{10}, \frac{2}{10}, \frac{4}{10})$	5.827	3.022	5.961	5.185	5.769	2.983	5.943	5.132

4.6. Real Data

Given that we do not have actual prior knowledge of the parameters, we decided to consider values around 2 for β and around $1/2$ for α, having only illustrative meaning. In Kundu's prior π_1, we take $a = 2, b = 1, c = 1, d = 2$. For Soland's prior π_2 we consider β uniformly distributed on $\{1.5, 1.75, 2, 2.25, 2.5\}$ and in order to determine the values of (a_i, b_i), we apply the method described before Lemma 1, if possible. Recall that in order to apply the method on β_i, there must exist a pair of records $r_j < r_l$ such that $lr_j^{\beta_i} < jr_l^{\beta_i}$. For $\beta_1 = 1.5$ and $\beta_2 = 1.75$, there exists no such pair of records while, for $\beta_3 = 2, \beta_4 = 2.25$ and $\beta_5 = 2.5$, the method can be applied and yields the following hyperparameter pairs (after rounding up to the nearest integer): $(40, 20)$, $(18, 8)$ and $(13, 5)$ for β_3, β_4 and β_5 respectively. For β_1 and β_2, where the method fails, we pick the value of the closest β, that is, $(40, 20)$. For numerical convenience, we analyze rainfall data using decimeters instead of millimeters, so that the values of δ and $\hat{r}_i$ are now divided by 100.

The results are shown in Tables 11 and 12. In both tables we observe that the estimates of the parameters using δ-records, with $\delta = -0.75$, outperform those based on records only, because they are closer to results using the complete sample. As in maximum likelihood prediction, the gain in prediction of future records using δ-records versus records is not clear; while there is some improvement using Soland's prior, this is not the case for Kundu's prior. This can be due to the surprising closeness of

Kundu's prediction using records to the prediction using the whole sample, which we believe happens by chance in this particular instance.

Table 11. Bayes estimates $\hat{\alpha}_B, \hat{\beta}_B$, Bayes predictions $\hat{R}_6, \hat{R}_7, \hat{R}_8$ and HPD intervals, for rainfall data, using Kundu's prior π_1.

δ	$\hat{\alpha}_B$	HPD α	$\hat{\beta}_B$	HPD β	$\hat{R}_6$	HPD R_6	$\hat{R}_7$	HPD R_7	$\hat{R}_8$	HPD R_8
0	0.41	[0.01, 1.05]	2.69	[1.21, 4.32]	3.04	[2.79, 3.61]	3.26	[2.79, 4.12]	3.47	[2.81, 4.58]
−0.25	0.41	[0.01, 1.04]	2.46	[1.06, 4.02]	3.13	[2.79, 3.87]	3.42	[2.79, 4.54]	3.24	[2.80, 3.84]
−0.50	0.39	[0.02, 0.97]	2.57	[1.19, 4.09]	3.07	[2.79, 3.67]	3.32	[2.79, 4.20]	3.55	[2.83, 4.68]
−0.75	0.48	[0.06, 1.05]	2.60	[1.44, 3.82]	2.98	[2.79, 3.37]	3.15	[2.79, 3.70]	3.32	[2.82, 4.00]
$-\infty$	0.58	[0.42, 0.74]	2.10	[1.72, 2.46]	3.05	[2.79, 3.54]	3.28	[2.80, 3.94]	3.50	[2.87, 4.30]

Table 12. Bayes estimation $\hat{\alpha}_B$, maximum probability estimation $\hat{\beta}_{\mathrm{MP}}$, Bayes predictions $\hat{R}_6, \hat{R}_7, \hat{R}_8$ and HPD intervals (except β), for rainfall data, using Soland's prior.

δ	$\hat{\alpha}_B$	HPD α	$\hat{\beta}_{\mathrm{MP}}$	$\hat{R}_6$	HPD R_6	$\hat{R}_7$	HPD R_7	$\hat{R}_8$	HPD R_8
0	0.39	[0.16, 0.62]	2.50	3.01	[2.79, 3.45]	3.21	[2.8, 3.79]	3.39	[2.85, 4.08]
−0.25	0.33	[0.15, 0.53]	2.50	3.04	[2.79, 3.53]	3.26	[2.8, 3.90]	3.47	[2.86, 4.22]
−0.50	0.34	[0.17, 0.52]	2.50	3.03	[2.79, 3.49]	3.24	[2.8, 3.84]	3.44	[2.86, 4.14]
−0.75	0.43	[0.27, 0.61]	2.50	2.98	[2.79, 3.34]	3.15	[2.8, 3.62]	3.31	[2.85, 3.86]
$-\infty$	0.47	[0.37, 0.56]	2.50	2.96	[2.79, 3.29]	3.12	[2.8, 3.54]	3.27	[2.85, 3.75]

5. Final Comments

Professionals interested in statistical inferences for the Weibull model, based on record data, should consider the possibility of using δ-records. We have presented in this paper the mathematical bases of the methodology, together with some theoretical results, such as consistency. We have also established the existence of estimators and predictors and discussed their numerical implementation so that researchers interested in applications can readily test the method on their own data. As commented above, we insist here that δ-records can be collected, in many cases, by slightly modifying the experimental setup for records so that the additional cost related to extra data is kept low. The conclusions to be drawn from the simulations and the application to the rainfall data are that δ-records improve inferences in the Weibull model. The impact is more notorious in MLE than in the Bayesian framework, possibly because of a strong influence of the priors. We believe that more investigation would be welcome for fine tuning this novel methodology. For example, it would be useful to have guidelines for the choice of δ and to explore the possibility of letting δ vary during data collection. Also, in the theoretical analysis of the model, it would be of interest to extend the consistency of Theorem 1 to the case of unknown shape parameter and study eventual asymptotic distributions; see Remark 4. These and other related topics will be considered in forthcoming papers.

6. Technical Results and Proofs

6.1. Consistency of the MLE $\hat{\lambda}(\beta)$

The proof of Theorem 1 follows from Lemmas 2–7, related to the asymptotic behavior of δ-records, which can be of independent interest. Strong convergence, as the appropriate index (usually n) tends to infinity, is denoted by $\xrightarrow{a.s.}$; inequalities with random variables are understood in the almost sure sense. Recall that R_m denotes the m-th record, while S_m and $(Y_{m,1}, \ldots, Y_{m,S_m})$ denote respectively the number and the vector of near-records associated to R_m.

Lemma 2. *If $\beta > 1$ then $S_n/n^\gamma \xrightarrow{a.s.} \infty$, for all $\gamma > 0$.*

Proof. It suffices to show that $S_n/\phi_{n,k} \xrightarrow{a.s.} \infty$, for $k \geq 1$, where $\phi_{n,k} = (n-1)\cdots(n-k)$ is the k-th falling factorial of $n-1$. To that end we apply the Borel-Cantelli lemma to show that $P(S_n < \phi_{n,k}M \text{ i.o.}) = 0$, for every integer $M > 0$ (i.o. stands for infinitely often). Observe that S_n conditional on R_n is geometrically distributed (starting at 0), with parameter

$$p_n := \frac{\bar{F}_{\lambda,\beta}(R_n)}{\bar{F}_{\lambda,\beta}(R_n+\delta)} = e^{-\alpha(R_n^\beta - (R_n+\delta)_+^\beta)}, \tag{26}$$

where $\alpha = \lambda^{-\beta}$, that is, $P(S_n = k|R_n) = (1-p_n)^k p_n, k \geq 0$; see Reference [10]. Noting that

$$R_n^\beta - (R_n+\delta)_+^\beta \geq \beta(R_n - (R_n+\delta)_+)(R_n+\delta)_+^{\beta-1} = |\delta|\beta(R_n+\delta)_+^{\beta-1},$$

from (26) we have $p_n \leq e^{-\alpha|\delta|\beta(R_n+\delta)_+^{\beta-1}}$ and so,

$$P(S_n < \phi_{n,k}M|R_n) = 1-(1-p_n)^{\phi_{n,k}M} \leq \phi_{n,k}Mp_n \leq \phi_{n,k}Me^{-\alpha|\delta|\beta(R_n+\delta)_+^{\beta-1}}.$$

Now, since R_n^β has Gamma(α, n) distribution (see Reference [2]), we get

$$P(S_n < \phi_{n,k}M) \leq M\int_0^\infty e^{-\alpha|\delta|\beta(t^{1/\beta}+\delta)_+^{\beta-1}} \phi_{n,k}\frac{\alpha^n t^{n-1}e^{-\alpha t}}{\Gamma(n)}dt.$$

Hence, knowing that the k-th factorial moment of a Poisson random variable with parameter μ is μ^k, we have

$$\begin{aligned}
\sum_{n=1}^\infty P(S_n < \phi_{n,k}M) &\leq M\int_0^\infty e^{-\alpha|\delta|\beta(t^{1/\beta}+\delta)_+^{\beta-1}} \alpha\sum_{n=1}^\infty \phi_{n,k}\frac{(\alpha t)^{n-1}e^{-\alpha t}}{\Gamma(n)}dt \\
&= M\int_0^\infty \alpha(\alpha t)^k e^{-\alpha|\delta|\beta(t^{1/\beta}+\delta)_+^{\beta-1}}dt \\
&= M\int_0^\infty \alpha(\alpha u^\beta)^k e^{-\alpha|\delta|\beta(u+\delta)_+^{\beta-1}}\beta u^{\beta-1}du < \infty,
\end{aligned}$$

and the conclusion follows. □

Lemma 3. *Let $U_n = \sum_{j=1}^{S_n}(Y_{n,j}^\beta - (R_n+\delta)_+^\beta)$, $n \geq 1$. Then, if $\beta > 1$, $U_n/S_n \xrightarrow{a.s.} \lambda^\beta$.*

Proof. From Proposition 1 in Reference [10] we know that, conditional on R_n and $S_n, S_n > 0$, the random variables $Y_{n,1}, \ldots, Y_{n,S_n}$ are iid, with common density function

$$g_{\lambda,\beta}(y) := \frac{\alpha\beta y^{\beta-1}e^{-\alpha y^\beta}}{e^{-\alpha(R_n+\delta)_+^\beta} - e^{-\alpha R_n^\beta}}, \qquad y \in ((R_n+\delta)_+, R_n),\ n \geq 1,$$

with $\alpha = \lambda^{-\beta}$. Hence, letting $\mathcal{F}$ be the σ-algebra generated by the sequences (R_n) and (S_n), we obtain

$$\begin{aligned}
E[Y_{n,j}^\beta|\mathcal{F}] &= \frac{1}{e^{-\alpha(R_n+\delta)_+^\beta} - e^{-\alpha R_n^\beta}}\int_{(R_n+\delta)_+}^{R_n} \alpha\beta y^{2\beta-1}e^{-\alpha y^\beta}dy \\
&= \frac{1}{e^{-\alpha(R_n+\delta)_+^\beta} - e^{-\alpha R_n^\beta}}\int_{(R_n+\delta)_+^\beta}^{R_n^\beta} \alpha t e^{-\alpha t}dt \\
&= \frac{1}{\alpha} - \frac{R_n^\beta e^{-\alpha R_n^\beta} - (R_n+\delta)_+^\beta e^{-\alpha(R_n+\delta)_+^\beta}}{e^{-\alpha(R_n+\delta)_+^\beta} - e^{-\alpha R_n^\beta}},
\end{aligned}$$

and so, $E[Y_{n,j}^{\beta} - (R_n+\delta)_+^{\beta}|\mathcal{F}] = \frac{1}{\alpha} - K_n$, where $K_n := \frac{R_n^{\beta}-(R_n+\delta)_+^{\beta}}{e^{\alpha(R_n^{\beta}-(R_n+\delta)_+^{\beta})}-1}$, $j = 1,\ldots,S_n$. Let also

$$Z_{n,j} = Y_{n,j}^{\beta} - E[Y_{n,j}^{\beta}|\mathcal{F}] = Y_{n,j}^{\beta} - (R_n+\delta)_+^{\beta} - \frac{1}{\alpha} + K_n,\ j = 1,\ldots,S_n, \tag{27}$$

and let $V_n = \sum_{j=1}^{S_n} Z_{n,j}, n \geq 1$. Then, by Tchebychev's inequality, $P(|V_n| > \epsilon S_n|\mathcal{F}) \leq \frac{\sigma_n^2}{\epsilon^2 S_n}$, for all $\epsilon > 0$, where

$$\sigma_n^2 := E(Z_{n,j}^2|\mathcal{F}) = E((Y_{n,j}^{\beta} - E(Y_{n,j}^{\beta}|\mathcal{F}))^2|\mathcal{F}) \leq E(Y_{n,j}^{2\beta}|\mathcal{F}) \leq R_n^{2\beta},\ n \geq 1.$$

Furthermore, since the sequence (R_n^{β}) is distributed as the ordered points of a homogeneous Poisson process, with rate $\lambda^{-\beta}$ (see Reference [2]), the strong law of large numbers implies

$$\frac{R_n^{\beta}}{n} \xrightarrow{a.s.} \lambda^{\beta}, \tag{28}$$

and so $\limsup \sigma_n^2/n^2 \leq \lambda^{2\beta}$. Then, by Lemma 2, with $\gamma = 4$, we have $\sigma_n^2/S_n = o(n^{-2})$. Therefore $\sum_{n=1}^{\infty} P(|V_n| > \epsilon S_n|\mathcal{F}) < \infty$ and, by the (conditional) Borel-Cantelli lemma, $P(|V_n| > \epsilon S_n \text{ i.o.}|\mathcal{F}) = 0$ and so $P(V_n/S_n \to 0|\mathcal{F}) = 1$. Then, taking expectation, we obtain $V_n/S_n \xrightarrow{a.s.} 0$. Finally, since $R_n^{\beta} - (R_n+\delta)_+^{\beta} \geq |\delta|\beta(R_n+\delta)_+^{\beta-1} \xrightarrow{a.s.} \infty$ (because $R_n \xrightarrow{a.s.} \infty$ and $\beta > 1$), we get $K_n \xrightarrow{a.s.} 0$. The conclusion now follows from the above convergence results and the identity $U_n/S_n = V_n/S_n + \lambda^{\beta} - K_n$. □

Lemma 4. *If $\beta = 1$ then $\frac{1}{n}\sum_{i=1}^{n} S_i \xrightarrow{a.s.} e^{-\delta/\lambda} - 1$.*

Proof. As stated in Lemma 2, conditionally on (R_n), the random variables S_i are independent and geometrically distributed, with parameter $p_i = e^{-\alpha(R_i-(R_i+\delta)_+)}$, where $\alpha = \lambda^{-1}$. Let $\tilde{S}_i = S_i - (1-p_i)/p_i, i \geq 1$, then $E(\tilde{S}_i|\mathcal{G}) = 0$, where $\mathcal{G}$ is the σ-algebra generated by (R_n). Furthermore, observe that $p_i \geq e^{\delta\alpha}$ and $E(\tilde{S}_i^4|\mathcal{G}) \leq 19p_i^{-4}$. So, the conditional fourth moment of $\tilde{S}_i$ is bounded above by $19e^{-4\alpha\delta}$ and the following (conditional) strong law of large numbers holds: $P\left(\frac{1}{n}\sum_{i=1}^{n}\tilde{S}_i \to 0 \middle| \mathcal{G}\right) = 1$. Then, taking expectation and observing that $(1-p_n)/p_n = e^{\alpha(R_n-(R_n+\delta)_+)} - 1 \xrightarrow{a.s.} e^{-\alpha\delta} - 1$, the conclusion follows. □

Lemma 5. *If $\beta = 1$ then $\frac{1}{n}\sum_{i=1}^{n} U_i \xrightarrow{a.s.} \lambda(e^{-\delta/\lambda} - 1) + \delta$, with U_i defined in Lemma 3.*

Proof. Let $\mathcal{F}, Z_{i,j}, V_i$ and α be defined as in the proof of Lemma 3. Note that, conditionally on $\mathcal{F}$, the $Z_{i,j}$ are independent with mean 0 and also that $|Z_{i,j}| \leq |\delta|$. It follows that $E(V_i^2|\mathcal{F}) \leq MS_i$ and $E(V_i^4|\mathcal{F}) \leq MS_i^2$, for some non-random constant $M > 0$. Consequently, since the V_i are also conditionally independent and centered, we have

$$E\left(\left(\sum_{i=1}^{n} V_i\right)^4 \middle| \mathcal{F}\right) = \sum_{i=1}^{n} E(V_i^4|\mathcal{F}) + 6\sum_{i<j} E(V_i^2|\mathcal{F})E(V_j^2|\mathcal{F}) \leq M\sum_{i=1}^{n} S_i^2 + 6M^2\sum_{i<j} S_iS_j.$$

On the other hand, from (26) and $R_i - (R_i+\delta)_+ \leq -\delta$, we get

$$E(S_i^2|R_i) \leq \frac{2}{p_i^2} \leq 2e^{-2\alpha\delta}, \quad E(S_iS_j|R_i,R_j) = E(S_i|R_i)E(S_j|R_j) \leq \frac{1}{p_ip_j} \leq e^{-2\alpha\delta}. \tag{29}$$

Then, by Tchebychev's inequality and (29),

$$P\left(\left|\frac{1}{n}\sum_{i=1}^{n}V_i\right|>\epsilon\right)\leq\frac{E((\sum_{i=1}^{n}V_i)^4)}{n^4\epsilon^4}$$
$$\leq\frac{M}{n^4\epsilon^4}\left(\sum_{i=1}^{n}E(S_i^2)+6M\sum_{i<j}E(S_iS_j)\right)$$
$$\leq\frac{M}{n^4\epsilon^4}(2ne^{-2\alpha\delta}+3n^2Me^{-2\alpha\delta}),$$

and it follows that $\frac{1}{n}\sum_{i=1}^{n}V_i\xrightarrow{a.s.}0$. The conclusion is obtained if we show that

$$\frac{1}{n}\sum_{i=1}^{n}S_i\left(\frac{1}{\alpha}-\frac{R_i-(R_i+\delta)_+}{e^{\alpha(R_i-(R_i+\delta)_+)}-1}\right)\xrightarrow{a.s.}\frac{1}{\alpha}(e^{-\alpha\delta}-1)+\delta,$$

but such convergence is implied by Lemma 4 and $R_n-(R_n+\delta)_+\xrightarrow{a.s.}-\delta$. $\square$

Lemma 6. *If $\beta<1$ then $\frac{1}{n}\sum_{i=1}^{n}S_i\xrightarrow{a.s.}0$.*

Proof. Let N_n and D_n be the number of records and of near-records among the first n observations, respectively, then it is clear that $S_1+\cdots+S_{N_n-1}\leq D_n\leq S_1+\cdots+S_{N_n}$. From Proposition 5.1 of Reference [28], we have $D_n/\log n\xrightarrow{a.s.}0$ and, noting that $N_{L(n)}=n$, where $L(n)$ is the n-th record time, we have

$$\frac{S_1+\cdots+S_{n-1}}{n}=\frac{S_1+\cdots S_{N_{L(n)}-1}}{N_{L(n)}}\leq\frac{D_{L(n)}}{\log L(n)}\frac{\log L(n)}{N_{L(n)}}\xrightarrow{a.s.}0,$$

where we have used $L(n)\xrightarrow{a.s.}\infty$ and the well-known result $N_n/\log n\xrightarrow{a.s.}1$, thus proving the result. $\square$

Lemma 7. *If $\beta<1$ then $\frac{1}{n}\sum_{i=1}^{n}U_i\xrightarrow{a.s.}0$, where U_i is defined in Lemma 3.*

Proof. Since $\beta<1$, we have $Y_{i,j}^{\beta}-(R_i+\delta)_+^{\beta}\leq R_i^{\beta}-(R_i+\delta)_+^{\beta}\leq|\delta|^{\beta}$. Therefore, by Lemma 6, $\frac{1}{n}\sum_{i=1}^{n}U_i\leq\frac{|\delta|^{\beta}}{n}\sum_{i=1}^{n}S_i\xrightarrow{a.s.}0$. $\square$

Proof of Theorem 1.
We divide the proof in three cases, depending on the tail behavior of $F_{\lambda,\beta}$, that is, $\beta>1$, $\beta=1$, $\beta<1$. Recall that $U_i=\sum_{j=1}^{S_i}(Y_{i,j}^{\beta}-(R_i+\delta)_+^{\beta})$, then from (6) and (7),

$$\hat{\lambda}(\beta)=\left(\frac{\sum_{i=1}^{n}\left(R_i^{\beta}-(R_i+\delta)_+^{\beta}+U_i\right)+R_n^{\beta}}{n+\sum_{i=1}^{n}S_i}\right)^{\frac{1}{\beta}}.$$

(i) Case $\beta>1$. Note that, by Lemma 2 and (28),

$$\frac{\sum_{i=1}^{n}\left(R_i^{\beta}-(R_i+\delta)_+^{\beta}\right)+R_n^{\beta}}{n+\sum_{i=1}^{n}S_i}\leq\frac{(n+1)R_n^{\beta}}{S_n}\xrightarrow{a.s.}0.$$

Then convergence of $\hat{\lambda}(\beta)$ to λ is equivalent to

$$\frac{\sum_{i=1}^{n}U_i}{n+\sum_{i=1}^{n}S_i}\xrightarrow{a.s.}\lambda^{\beta},$$

which follows from Lemmas 2 and 3.

(ii) Case $\beta = 1$. From Lemma 4 we have $\frac{N}{n} = \frac{n+\sum_{i=1}^{n} S_i}{n} \xrightarrow{a.s.} e^{-\delta/\lambda}$. Also, since $R_n \xrightarrow{a.s.} \infty$, we have $\frac{1}{n}\sum_{i=1}^{n}(R_i - (R_i+\delta)_+) \xrightarrow{a.s.} -\delta$ and the result follows from (28) and Lemma 5.
(iii) Case $\beta < 1$. Note that $R_n^\beta - (R_n+\delta)_+^\beta \xrightarrow{a.s.} 0$, since $R_n \xrightarrow{a.s.} \infty$ and $\beta < 1$, so $\frac{1}{n}\sum_{i=1}^{n}(R_i^\beta - (R_i+\delta)_+^\beta) \xrightarrow{a.s.} 0$. This convergence, together with (28) and Lemmas 6 and 7, complete the proof. □

6.2. MLP of R_m

Proof of Proposition 4.
(i) For $z > r_n$, let $l_1(z,\lambda,\beta) = \log L_1^P(z,\mathbf{t}|\lambda,\beta)$. Then

$$l_1(z,\lambda,\beta) = (m-n-1)\log(z^\beta - r_n^\beta) - \log\Gamma(m-n) - \beta(m-n+N)\log\lambda + (N+1)\log\beta \\ + (\beta-1)\log z - \lambda^{-\beta}(z^\beta - r_n^\beta + G(\beta)) + (\beta-1)\log J.$$

Solving for λ in

$$\frac{\partial l_1}{\partial\lambda} = -\frac{\beta(m-n+N)}{\lambda} + \beta\lambda^{-(\beta+1)}(z^\beta - r_n^\beta + G(\beta)) = 0,$$

we obtain the stationary point $\tilde{\lambda}(\beta) = \left(\frac{z^\beta - r_n^\beta + G(\beta)}{m-n+N}\right)^{\frac{1}{\beta}} > 0$, which is the unique element of $\underset{\lambda}{\operatorname{argmax}}\, l_1(z,\lambda,\beta)$, because $\frac{\partial^2 l_1}{\partial\lambda^2} < 0$ at $\tilde{\lambda}$, for all z,β. Thus, the problem $\max_{\lambda,z} l_1(z,\lambda,\beta)$ is reduced to $\max_z l_1(z,\tilde{\lambda}(\beta),\beta)$ or, equivalently, to $\max_z \tilde{l}_1(z)$, where

$$\tilde{l}_1(z) := -(m-n+N)\log(z^\beta - r_n^\beta + G(\beta)) + (N+1)\log\beta \\ + (m-n-1)\log(z^\beta - r_n^\beta) + (\beta-1)\log z + \beta\log J.$$

With the change of variable $x = z^\beta - r_n^\beta$, we get

$$\hat{l}_1(x) := -(m-n+N)\log(x+G(\beta)) + (N+1)\log\beta + (m-n-1)\log x + \tfrac{\beta-1}{\beta}\log(x+r_n^\beta) + \beta\log J, \tag{30}$$

and $\max_{z>r_n} \tilde{l}_1(z)$ is reduced to $\max_{x>0} \hat{l}_1(x)$. We now study the sign of the derivative

$$\frac{\partial\hat{l}_1}{\partial x} = -\frac{N+m-n}{x+G(\beta)} + \frac{m-n-1}{x} + \frac{\beta-1}{\beta(x+r_n^\beta)}, \tag{31}$$

which is equal to the sign of the polynomial $q(x) := ax^2 + bx + c$, with coefficients defined in (10).

We observe, since $a < 0$ and $ac \le 0$, that q is concave on $\mathbb{R}$ and has real roots given by $\nu = \frac{-b+\sqrt{b^2-4ac}}{2a} \le 0$ and $\rho = \frac{-b-\sqrt{b^2-4ac}}{2a} \ge 0$. If $m = n+1$, then $c = 0$ and $b = (\beta-1)G(\beta) - \beta(N+1)r_n^\beta < 0$, because $G(\beta) \le (N+1)r_n^\beta$; see (6), for definitions of N and $G(\beta)$. Hence $\rho = -\frac{b+|b|}{2a} = 0$. On the other hand, if $m > n+1$, then $c > 0$ and $ac < 0$, which clearly implies $\rho > 0$. We summarize the above in terms of (31), denoting by $\tilde{z}(\beta)$ an element of $\underset{z}{\operatorname{argmax}}\, \tilde{l}_1(z)$:

- If $m = n+1$, then $\rho = 0$, $\frac{\partial\hat{l}_1}{\partial x}$ is negative on $(0,+\infty)$ and so, $\hat{l}_1(x)$ has no maximum on $(0,+\infty)$. In this case we do, however, define $\tilde{z}(\beta) = r_n$, that is, $\tilde{z}^\beta(\beta) = \rho + r_n^\beta$.
- If $m > n+1$, then $\rho > 0$, $\frac{\partial\hat{l}_1}{\partial x}$ is positive on $(0,\rho)$ and is negative on $(\rho,+\infty)$. Hence $\tilde{z}^\beta(\beta) = \rho + r_n^\beta$.

(ii) Given the definitions of $\tilde{z}(\beta)$ and $\tilde{\lambda}(\beta)$, it only remains to prove that $\tilde{A}$ is nonempty, noting that $\tilde{A} = \underset{\beta}{\operatorname{argmax}}\, \hat{l}_1(\rho)$. By Lemma 8, $\hat{l}_1(\rho) \to -\infty$, as $\beta \to 0^+$ and, by Lemma 10, $\hat{l}_1(\rho) \to -\infty$, as $\beta \to +\infty$. This ensures that $\tilde{A}$ is nonempty, since $\hat{l}_1(\rho)$ is a continuous function of $\beta > 0$. □

Lemma 8. *(i)* $\lim_{\beta\to 0^+} \frac{\rho}{\beta} \to m-n-1$,

(ii) $\lim_{\beta\to 0^+} \hat{l}_1(\rho) \to -\infty$.

Proof. (i) Note that $r_n^\beta \to 1$ and $G(\beta) \to 1$, implying $a \to -1, b \to -1, c \to 0$ and $c/\beta \to m-n-1$, as $\beta \to 0^+$. Hence, the conclusion follows from

$$\frac{\rho}{\beta} = \frac{2c}{\beta\left(-b+\sqrt{b^2-4ac}\right)} \to m-n-1. \tag{32}$$

(ii) Let $m = n+1$ and recall that $\rho = 0$. Then the conclusion is reached, since, as $\beta \to 0^+$,

$$\hat{l}_1(\rho) = (N+1)(\log\beta - \log G(\beta)) + \beta(\log r_n + \log J) - \log r_n \to -\infty.$$

On the other hand, if $m > n+1$, from $G(\beta) \to 1, \rho \to 0$, we have

$$\lim_{\beta\to 0^+} \hat{l}_1(\rho) = \lim_{\beta\to 0^+}((N+1)\log\beta + (m-n-1)\log\rho) + \lim_{\beta\to 0^+} \frac{\beta-1}{\beta}\log(\rho + r_n^\beta). \tag{33}$$

The first limit in the RHS of (33) is $-\infty$ and the second is finite and can be computed from (32), since

$$(\rho + r_n^\beta)^{1/\beta} = \left(1+\beta\left(\frac{\rho}{\beta} + \frac{r_n^\beta - 1}{\beta}\right)\right)^{1/\beta} \to r_n e^{m-n-1}.$$

□

Lemma 9. *Let $m > n+1$. Then there exist constants $A, B > 0$ such that, for all $\beta > 1$,*

$$A r_n^\beta < \rho < B r_n^\beta. \tag{34}$$

Proof. Note that $G(\beta) = 2r_n^\beta + o(r_n^\beta)$, as $\beta \to +\infty$. Then, from (10), $a/\beta = -N + o(1)$, $b/\beta = r_n^\beta(K + o(1))$ and $c/\beta = 2(m-n-1)r_n^{2\beta}(1+o(1))$, as $\beta \to +\infty$, where $K \neq 0$ is a constant. Then

$$\frac{\rho}{r_n^\beta} \to \frac{K + \sqrt{K^2 + 8N(m-n-1)}}{2N} > 0,$$

and the conclusion follows. □

Lemma 10. $\lim_{\beta\to+\infty} \hat{l}_1(\rho) = -\infty$.

Proof. Let $m > n+1$. From Lemma 9 and recalling that $r_n^\beta < G(\beta) \le (N+1)r_n^\beta$, we have

$$\begin{aligned}\hat{l}_1(\rho) &\le -(m-n+N)\log((A+1)r_n^\beta) + (N+1)\log\beta \\ &\quad + (m-n-1)\log(Br_n^\beta) + \frac{\beta-1}{\beta}\log((B+1)r_n^\beta) + \beta\log J \\ &= \beta(\log J - N\log r_n) + (N+1)\log\beta + K + o(1) \to -\infty,\end{aligned}$$

as $\beta \to +\infty$, where K is a constant and the limit is a consequence of the inequality $\log J < N \log r_n$; see (6) for definitions of J and N. Now, if $m = n+1$, we have $\rho = 0$ and

$$\begin{aligned}\hat{l}_1(\rho) &= (N+1)(\log\beta - \log G(\beta)) + (\beta-1)\log r_n + \beta\log J \\ &\leq (N+1)(\log\beta - \beta\log r_n) + \beta(\log r_n + \log J) - \log r_n \\ &= (N+1)\log\beta + \beta(\log J - N\log r_n) - \log r_n \to -\infty.\end{aligned}$$

□

Remark 7. *It is clear that the maximization of $\hat{l}_1(\rho)$ with respect to β, in the proof of Proposition 4, cannot be done analytically and numerical techniques must be used. The search for the maximum can be reduced to a compact interval $[\beta_L, \beta_U]$, with $\beta_L > 0$, as explained below.*

Let us write $\rho(\beta)$ instead of just ρ, to make explicit the dependence on β. Then, by Lemma 8 (ii), there exists $\beta_L \in (0,1)$ such that $\hat{l}_1(\rho(\beta)) < \hat{l}_1(\rho(1))$, for all $\beta < \beta_L$ and hence $\operatorname{argmax}_\beta \hat{l}_1(\rho(\beta)) \subset [\beta_L, \infty)$. For the upper bound β_U we observe, from Lemma 10, that a value $\beta_U > 1$ can be found, such that $\hat{l}_1(\rho(\beta)) < \hat{l}_1(\rho(1))$, for all $\beta > \beta_U$. Therefore, $\operatorname{argmax}_\beta \hat{l}_1(\rho(\beta)) \subset (-\infty, \beta_U]$ and, since $\beta_L < \beta_U$, we have $\operatorname{argmax}_\beta \hat{l}_1(\rho(\beta)) \subset [\beta_L, \beta_U]$. Note that, when writing $\hat{l}_1(\rho(1))$, we mean that β is replaced by 1 in the whole expression for $\hat{l}_1$, not only in $\rho(\beta)$.

Author Contributions: All the authors contributed significantly to the present paper. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by project PIA AFB-170001, Fondecyt grant 1161319 and project MTM2017-83812-P of MINECO. The third author gratefully acknowledges support from a *Gran Mariscal de Ayacucho Foundation* grant. The first, second and fourth authors are members of the research group Modelos Estocásticos of DGA.

Acknowledgments: We thank the anonymous reviewers for their careful reading of the manuscript and their useful comments and suggestions, which improved the presentation of the paper.

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

References

1. Ahsanullah, M.; Nevzorov, V.B. *Records via Probability Theory*; Springer: Berlin/Heidelberg, Germany, 2015.
2. Arnold, B.; Balakrishnan, N.; Nagaraja, H. *Records*; John Wiley & Sons: New York, NY, USA, 1998.
3. Gulati, D.; Padgett, W. *Parametric and Nonparametric Inference from Record-Breaking Data*; *Lecture Notes in Statistics*; Springer: New York, NY, USA, 2003; Volume 72.
4. Soliman, A.; Abd Ellah, A.; Sultan, S. Comparison of estimates using record statistics from Weibull model: Bayesian and non-Bayesian approaches. *Comput. Stat. Data Anal.* **2006**, *51*, 2065–2077. [CrossRef]
5. Wang, B.X.; Ye, Z.S. Inference on the Weibull distribution based on record values. *Comput. Stat. Data Anal.* **2015**, *83*, 26–36. [CrossRef]
6. Khmaladze, E.; Nadareishvili, M.; Nikabadze, A. Asymptotic behaviour of a number of repeated records. *Stat. Probab. Lett.* **1997**, *35*, 49–58. [CrossRef]
7. Balakrishnan, N.; Pakes, A.; Stepanov, A. On the number and sum of near-record observations. *Adv. Appl. Probabab.* **2005**, *37*, 765–780. [CrossRef]
8. Castaño-Martínez, A.; López-Blázquez, F.; Salamanca-Miño, B. Exceedances of records. *Metrika* **2016**, *79*, 837–866. [CrossRef]
9. Gouet, R.; López, F.J.; Sanz, G. Asymptotic normality for the counting process of weak records and δ-records in discrete models. *Bernoulli* **2007**, *13*, 754–781. [CrossRef]
10. Gouet, R.; López, F.J.; Sanz, G. On δ-record observations: Asymptotic rates for the counting process and elements of maximum likelihood estimation. *TEST* **2012**, *21*, 188–214. [CrossRef]
11. Gouet, R.; Lafuente, M.; López, F.J.; Sanz, G. δ-Records Observations in Models with Random Trend. In *The Mathematics of the Uncertain: A Tribute to Pedro Gil*; Gil, E., Gil, E., Gil, J., Gil, M.Á., Eds.; Springer International Publishing: Cham, Switzerland, 2018; pp. 209–217.
12. López-Blázquez, F.; Salamanca-Miño, B. Distribution theory of δ-record values. Case $\delta \leq 0$. *TEST* **2015**, *24*, 715–738. [CrossRef]

13. López-Blázquez, F.; Salamanca-Miño, B. Distribution theory of δ-record values. Case $\delta \geq 0$. *TEST* **2015**, *24*, 558–582. [CrossRef]
14. Wergen, G. Records in stochastic processes—Theory and applications. *J. Phys. A Math. Theor.* **2013**, *46*, 223001. [CrossRef]
15. Gouet, R.; López, F.J.; Maldonado, L.; Sanz, G. Statistical inference for the geometric distribution based on δ-records. *Comput. Stat. Data Anal.* **2014**, *78*, 21–32. [CrossRef]
16. Shahzad, M.; Ullah, E.; Hussanan, A. Beta Exponentiated Modified Weibull Distribution: Properties and Application. *Symmetry* **2019**, *11*, 781. [CrossRef]
17. Rao, G.; Albassam, M.; Aslam, M. Evaluation of Bootstrap Confidence Intervals Using a New Non-Normal Process Capability Index. *Symmetry* **2019**, *11*, 484. [CrossRef]
18. Dey, S.; Dey, T.; Luckett, D.J. Statistical Inference for the Generalized Inverted Exponential Distribution Based on Upper Record Values. *Math. Comput. Simul.* **2016**, *120*, 64–78. [CrossRef]
19. Jafari, A.A.; Zakerzadeh, H. Inference on the parameters of the Weibull distribution using records. *SORT* **2015**, *39*, 3–18.
20. Jaheen, Z. Empirical Bayes analysis of record statistics based on linex and quadratic loss functions. *Comput. Math. Appl.* **2004**, *47*, 947–954. [CrossRef]
21. Madi, M.T.; Raqab, M.Z. Bayesian prediction of rainfall records using the generalized exponential distribution. *Environmetrics* **2007**, *18*, 541–549. [CrossRef]
22. Teimouri, M.; Gupta, A.K. On the Weibull record statistics and associated inferences. *Statistica* **2012**, *72*, 145–162.
23. Teimouri, M.; Nadarajah, S. Bias corrected MLEs for the Weibull distribution based on records. *Stat. Methodol.* **2013**, *13*, 12–24. [CrossRef]
24. Wu, J.W.; Tseng, H.C. Statistical inference about the shape parameter of the Weibull distribution by upper record values. *Stat. Pap.* **2007**, *48*, 95–129. [CrossRef]
25. Kundu, D.; Howlader, H. Bayesian inference and prediction of the inverse Weibull distribution for Type-II censored data. *Comput. Stat. Data Anal.* **2010**, *54*, 1547–1558. [CrossRef]
26. Nordman, D.J.; Meeker, W.Q. Weibull Prediction Intervals for a Future Number of Failures. *Technometrics* **2002**, *44*, 15–23. [CrossRef]
27. Rinne, H. *The Weibull Distribution: A Handbook*; CRC Press: Boca Raton, FL, USA, 2008.
28. Gouet, R.; López, F.J.; Sanz, G. Limit theorems for the counting process of near-records. *Commun. Stat. Simul. Comput.* **2012**, *71*, 820–832. [CrossRef]
29. Gouet, R.; López, F.J.; Sanz, G. Central limit theorem for the number of near-records. *Commun. Stat.-Theory Methods* **2012**, *41*, 309–324. [CrossRef]
30. Basak, P.; Balakrishnan, N. Maximum likelihood prediction of future record statistic. In *Mathematical and Statistical Methods in Reliability*; Lindquist, B.B.H., Doksun, K., Eds.; World Scientific Publishing: Singapore, 2003; pp. 159–175.
31. Kundu, D. Bayesian Inference and Life Testing Plan for the Weibull Distribution in Presence of Progressive Censoring. *Technometrics* **2008**, *50*, 144–154. [CrossRef]
32. Soland, R.M. Bayesian Analysis of the Weibull Process With Unknown Scale and Shape Parameters. *IEEE Trans. Reliab.* **1969**, *R-18*, 181–184. [CrossRef]
33. Al-Hussaini, E.K. Predicting observables from a general class of distributions. *J. Stat. Plan. Inference* **1999**, *79*, 79–91. [CrossRef]
34. Ahmadi, J.; Doostparast, M. Bayesian estimation and prediction for some life distributions based on record values. *Stat. Pap.* **2006**, *47*, 373–392. [CrossRef]
35. Diaconis, P.; Freedman, D. On the consistency of Bayes estimates. *Ann. Stat.* **1986**, *14*, 1–26. [CrossRef]