

Empacher, Christina; Kamps, Udo

**Article — Published Version**

## Simultaneous prediction of record values

Metrika

**Provided in Cooperation with:**

Springer Nature

*Suggested Citation:* Empacher, Christina; Kamps, Udo (2025) : Simultaneous prediction of record values, *Metrika*, ISSN 1435-926X, Springer, Berlin, Heidelberg, Vol. 88, Iss. 6, pp. 1571-1589, <https://doi.org/10.1007/s00184-025-00998-0>

This Version is available at:

<https://hdl.handle.net/10419/330431>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<https://creativecommons.org/licenses/by/4.0/>



# Simultaneous prediction of record values

Christina Empacher<sup>1</sup> · Udo Kamps<sup>1</sup>

Received: 4 September 2024 / Accepted: 4 May 2025 / Published online: 21 May 2025  
© The Author(s) 2025

## Abstract

The prediction of record values based on a single sample has been studied extensively in the literature. The case of two observed samples of records has also been considered, where a record of one sample is predicted based on data of both samples. In a multi-sample set-up, simultaneous prediction of future record values in the samples is examined and prediction is based on information from observed records of all samples. Known point prediction procedures in the one-sample case, namely maximum likelihood, maximum observed likelihood and maximum product of spacings prediction, are adapted to the described setting of several samples. The point prediction methods are exemplified and elaborated by results for an underlying exponential and Pareto distribution. As an application, the point predictors are applied to data from athletics, where continental records from men's javelin, which may be assumed to be based on an underlying Pareto distribution, are simultaneously predicted.

**Keywords** Point prediction · Maximum product of spacings prediction · Multi-sample set-up · Exponential distribution · Pareto distribution · Real data analysis

## 1 Introduction

Methods of point and interval prediction of future record values based on observed record values in a single sequence of independent and identically distributed (iid) random variables have been studied in the literature in case of different distributions, frequently, and they can be applied in fields such as environmetrics, insurance mathematics and sports data analytics; we refer to, e.g., Basak and Balakrishnan (2003), Madi and Raqab (2004), Raqab (2007), Raqab et al (2007), Almuzaini and Raqab

---

C. Empacher, U. Kamps these authors contributed equally to this work.

---

✉ Christina Empacher  
christina.empacher@rwth-aachen.de

Udo Kamps  
udo.kamps@rwth-aachen.de

<sup>1</sup> Institute of Statistics, RWTH Aachen University, Aachen D-52056, Germany

(2022), Navarro (2022), Volovskiy and Kamps (2023), Empacher et al (2024), and the references therein.

There are also situations in applications, where we observe record values in independent samples and where we are interested in statistical inference based on all record information. Several articles examine the two-sample situation and consider prediction of future record values in one sample by additionally utilizing the realizations of record values in the second sample. Such procedures have also been developed for other models of ordered random variables or for combinations of such models. For relevant literature we refer to, e.g., Raqab and Balakrishnan (2008), Ahmadi and MirMostafaei (2009), Raqab (2009), Ahmadi et al (2010), Ahmadi and Balakrishnan (2010), Singh et al (2017), Basiri et al (2020), Almuzaini and Raqab (2022).

In this article, we study a multi-sample set-up of  $m \in \mathbb{N}$  independent samples, aiming at simultaneous point prediction of future record values in all samples based on observed record values in the  $m$  samples, where the underlying distributions of the iid random variables in the samples coincide.

First, we recall the one-sample record setting. Let  $X_1, X_2, \dots$  be a sequence of iid random variables with an absolutely continuous cumulative distribution function (cdf)  $F_\theta$ , probability density function (pdf)  $f_\theta$  and unknown  $d$ -dimensional parameter  $\theta \in \Theta \subset \mathbb{R}^d$ ,  $d \in \mathbb{N}$ . Let  $\alpha(F_\theta)$  and  $\omega(F_\theta)$  denote the left and right endpoint of the support of  $F_\theta$ . Then, record times are defined via

$$T(1) = 1, \quad T(n+1) = \min\{j > T(n) : X_j > X_{T(n)}\}, \quad n \in \mathbb{N}.$$

Then, the random variables  $R_n = X_{T(n)}$ ,  $n \in \mathbb{N}$ , are called (upper) record values or records. The joint density function of the first  $r$  records is given by

$$f_\theta^{R_1, \dots, R_r}(x_1, \dots, x_r) = f_\theta(x_r) \prod_{i=1}^{r-1} \frac{f_\theta(x_i)}{1 - F_\theta(x_i)}, \quad \alpha(F_\theta) < x_1 < \dots < x_r < \omega(F_\theta)$$

and their marginal density function by

$$f_\theta^{R_r}(x_r) = \frac{1}{(r-1)!} (-\ln(1 - F_\theta(x_r)))^{r-1} f_\theta(x_r), \quad x_r \in (\alpha(F_\theta), \omega(F_\theta)). \quad (1)$$

Let  $\mathbb{R}_{<}^r = \{(x_1, \dots, x_r) \in \mathbb{R}^r \mid \alpha(F_\theta) < x_1 < \dots < x_r < \omega(F_\theta)\}$  denote the set of strictly ordered observations in  $\mathbb{R}^r$ .

In the literature several procedures for the point and interval prediction of a future record  $R_s$  based on observed records  $R_1, \dots, R_r$  with  $r < s$  are studied (see, e.g., Volovskiy and Kamps (2020a), Navarro (2022), Empacher et al (2023) and the references therein). In this work we focus on likelihood based predictors, namely the maximum likelihood predictor (MLP), the maximum observed likelihood predictor (MOLP), as well as on the maximum product of spacings predictor (MPSP). The MLP of a future or unobserved random variable  $Y$  based on an observed random vector  $X$  is obtained by maximizing the joint density function  $f_\theta^{X,Y}$  of  $X$  and  $Y$  regarding  $y$  and the unknown parameter  $\theta \in \Theta$  of Kaminsky and Rhodin (1985).

Analogously, the MOLP is derived by maximizing the conditional density function  $f_\theta^{X|Y}$  of  $X$  given  $Y$  regarding  $y$  and  $\theta$  (see Volovskiy and Kamps (2020a)).

Based on record values  $R_1, \dots, R_r$  it is shown in Volovskiy and Kamps (2020a) that the MOLP of  $R_s, s > r$ , is given by

$$\pi_{MOLP}^{(s)}(R_1, \dots, R_r) = F_{\hat{\theta}_{MOL}}^{-1} \left( 1 - \left( 1 - F_{\hat{\theta}_{MOL}}(R_r) \right)^{\frac{s-1}{r}} \right)$$

where  $\hat{\theta}_{MOL} = \hat{\theta}_{MOL}(R_1, \dots, R_r)$  and  $\hat{\theta}_{MOL}(x_1, \dots, x_r)$  maximizes the function  $\Psi$  given by

$$\Psi(\theta \mid x_1, \dots, x_r) = \prod_{j=1}^r \frac{f_{\theta}(x_j)}{(1 - F_{\theta}(x_j))(-\ln(1 - F_{\theta}(x_r)))} \mathbb{I}_{(\alpha(F_{\theta}), \omega(F_{\theta}))^r}(x_1, \dots, x_r)$$

regarding  $\theta \in \Theta$ .

The MPSP of  $R_s, s > r$ , is obtained in Volovskiy and Kamps (2020b) by maximizing the function  $\mathcal{P}_s$  given by

$$\mathcal{P}_s(x_1, \dots, x_s, \theta) = \prod_{j=1}^s \left( \frac{H_{\theta}(x_j)}{H_{\theta}(x_s)} - \frac{H_{\theta}(x_{j-1})}{H_{\theta}(x_s)} \right)$$

regarding  $\theta \in \Theta$  and  $x_{r+1} < \dots < x_s$ , where  $H_{\theta}(x) = -\ln(1 - F_{\theta}(x))$  is the cumulative hazard rate at  $x$  with  $H_{\theta}(x_0) = 0$ . It turns out that the MPSP has the form

$$\pi_{MPSP}^{(s)}(R_1, \dots, R_r) = F_{\hat{\theta}_{MPS}}^{-1} \left( 1 - \left( 1 - F_{\hat{\theta}_{MPS}}(R_r) \right)^{s/r} \right)$$

where  $\hat{\theta}_{MPS} = \hat{\theta}_{MPS}(R_1, \dots, R_r)$  and  $\hat{\theta}_{MPS}(x_1, \dots, x_r)$  maximizes  $\mathcal{P}_r$  regarding  $\theta \in \Theta$ .

The forms of the predictors are utilized in Sect. 2.

Other prediction procedures such as linear prediction (Ahsanullah 1980) and prediction using quantile regression and distortion functions (Navarro 2022) are considered in the literature.

In the following section, representations of the MLP, MOLP and MPSP for the simultaneous prediction of records based on an arbitrary number  $m$  of iid samples are derived. These results may be used to construct a joint prediction interval for future records in the samples. The point prediction results are exemplified in Sect. 3 by specific and explicit results for an underlying exponential and Pareto distribution. Finally, in Sect. 4, the MPSP for Pareto records is applied to data from athletics, namely to series of continental records from men’s javelin.

## 2 Point prediction based on several samples

We consider record values from  $m$  independent samples and aim at predicting future records by means of MLP, MOLP as well as MPSP in every sample simultaneously.

Let  $X_j^{(i)} \stackrel{iid}{\sim} F_\theta, i = 1, \dots, m, j = 1, 2, \dots$ , be  $m$  independent sequences of random variables. Then,  $R_1^{(i)}, R_2^{(i)}, \dots$  denotes the sequence of upper record values in sample  $i$  based on  $X_1^{(i)}, X_2^{(i)}, \dots$ . The first  $r_i$  records in the  $i$ -th sequence are denoted by  $R_\star^{(i)} = (R_1^{(i)}, \dots, R_{r_i}^{(i)})$  which will be assumed to be observed. The future records  $R_{s_1}^{(1)}, \dots, R_{s_m}^{(m)}$  with  $s_1 > r_1, \dots, s_m > r_m$  for each sample of record values are to be predicted simultaneously.

Due to the independence of the sequences of record values, the respective predictive likelihood function is given by

$$\begin{aligned}
 PL(x_\star^{(1)}, x_{s_1}^{(1)}, \dots, x_\star^{(m)}, x_{s_m}^{(m)}, \theta) &= \prod_{i=1}^m f_\theta^{R_\star^{(i)}, R_{s_i}^{(i)}}(x_\star^{(i)}, x_{s_i}^{(i)}) \\
 &= \prod_{i=1}^m \left( \frac{1}{(s_i - r_i - 1)!} \left( \prod_{j=1}^{r_i} \frac{f_\theta(x_j^{(i)})}{1 - F_\theta(x_j^{(i)})} \right) f_\theta(x_{s_i}^{(i)}) \right. \\
 &\quad \left. \cdot \left( -\ln \left( \frac{1 - F_\theta(x_{s_i}^{(i)})}{1 - F_\theta(x_{r_i}^{(i)})} \right) \right)^{s_i - r_i - 1} \right). \tag{2}
 \end{aligned}$$

In order to obtain a MLP, this function has to be maximized regarding  $x_{s_1}^{(1)}, \dots, x_{s_m}^{(m)}$  and  $\theta$ . However, a general representation is not available.

Similar to the one-sample case examined in Volovskiy and Kamps (2020a, b), both the MOLP and the MPSP take a specific form as shown in Theorems 1 and 2, where the maximization procedure is reduced to a maximization regarding the distribution parameter  $\theta$ .

**Theorem 1 (MOLP)** For  $x_\star^{(i)} = (x_1^{(i)}, \dots, x_{r_i}^{(i)}) \in \mathbb{R}_{<}^{r_i}, 2 \leq r_i \in \mathbb{N}, i = 1, \dots, m$ , let the function  $\Psi$  be defined by

$$\begin{aligned}
 \Psi(\theta \mid x_\star^{(1)}, \dots, x_\star^{(m)}) &= \prod_{i=1}^m \left( \frac{\prod_{j=1}^{r_i} f_\theta(x_j^{(i)})}{\left( \prod_{j=1}^{r_i} (1 - F_\theta(x_j^{(i)})) \right) \left( -\ln(1 - F_\theta(x_{r_i}^{(i)})) \right)} \right) \\
 &\quad \cdot \mathbb{I}_{[\alpha(F_\theta), \omega(F_\theta)]^{r_i}}(x_\star^{(i)})
 \end{aligned}$$

for  $\theta \in \Theta$ . If, for  $x_\star \in \mathbb{R}_{<}^r$ , a Borel measurable function  $\widehat{\theta}: \mathbb{R}^{\sum_{i=1}^m r_i} \rightarrow \Theta$  exists with

$$\Psi(\widehat{\theta}(x_\star^{(1)}, \dots, x_\star^{(m)}) \mid x_\star^{(1)}, \dots, x_\star^{(m)}) = \max_{\theta \in \Theta} \Psi(\theta \mid x_\star^{(1)}, \dots, x_\star^{(m)}),$$

then for  $s_i \in \mathbb{N}$  with  $s_i \geq r_i + 1$  the MOLP of  $R_{s_i}^{(i)}$  is determined by

$$\pi_{MOLP}^{(i)} \left( R_{\star}^{(1)}, \dots, R_{\star}^{(m)} \right) = F_{\widehat{\theta}_{MOL}}^{-1} \left( 1 - \left( 1 - F_{\widehat{\theta}_{MOL}} \left( R_{r_i}^{(i)} \right) \right)^{\frac{s_i-1}{r_i}} \right)$$

for all  $i = 1, \dots, m$  simultaneously with

$$\widehat{\theta}_{MOL} = \widehat{\theta} \left( R_{\star}^{(1)}, \dots, R_{\star}^{(m)} \right).$$

If  $s_i = r_i + 1$  for some  $i \in \{1, \dots, m\}$ , the respective  $i$ -th MOLP is given by  $\pi_{MOLP}^{(i)} \left( R_{\star}^{(1)}, \dots, R_{\star}^{(m)} \right) = R_{r_i}^{(i)}$ .

**Proof** With the independence of the sequences of record values and Equations (1) and (2), the conditional density function of  $R_{\star}^{(1)}, \dots, R_{\star}^{(m)}$  given  $R_{s_1}^{(1)}, \dots, R_{s_m}^{(m)}$  is

$$\begin{aligned} & \frac{PL \left( x_{\star}^{(1)}, x_{s_1}^{(1)}, \dots, x_{\star}^{(m)}, x_{s_m}^{(m)}, \theta \right)}{\prod_{i=1}^m f_{\theta}^{R_{s_i}^{(i)}} \left( x_{s_i}^{(i)} \right)} \\ &= \prod_{i=1}^m \frac{(s_i - 1)!}{(s_i - r_i - 1)!} \left( \prod_{j=1}^{r_i} \frac{f_{\theta} \left( x_j^{(i)} \right)}{\left( 1 - F_{\theta} \left( x_j^{(i)} \right) \right) \left( -\ln \left( 1 - F_{\theta} \left( x_{r_i}^{(i)} \right) \right) \right)} \right) \\ & \bullet \left( 1 - \frac{\ln \left( 1 - F_{\theta} \left( x_{r_i}^{(i)} \right) \right)}{\ln \left( 1 - F_{\theta} \left( x_{s_i}^{(i)} \right) \right)} \right)^{s_i - r_i - 1} \left( \frac{\ln \left( 1 - F_{\theta} \left( x_{r_i}^{(i)} \right) \right)}{\ln \left( 1 - F_{\theta} \left( x_{s_i}^{(i)} \right) \right)} \right)^{r_i}. \end{aligned}$$

Because of its product structure this function can be maximized regarding  $x_{s_1}^{(1)}, \dots, x_{s_m}^{(m)}$  separately. Let, at first,  $s_i > r_i + 1$  for all  $i \in \{1, \dots, m\}$ . Then the last two factors take the shape of a beta density function. Hence, they are maximized by the mode of that beta distribution, which is given by

$$\frac{\ln \left( 1 - F_{\theta} \left( x_{r_i}^{(i)} \right) \right)}{\ln \left( 1 - F_{\theta} \left( x_{s_i}^{(i)} \right) \right)} = \frac{r_i}{s_i - 1}.$$

Therefore, the optimal  $x_{s_i}^{(i)}$  is determined as

$$x_{s_i}^{(i)} = F_{\theta}^{-1} \left( 1 - \left( 1 - F_{\theta} \left( x_{r_i} \right) \right)^{\frac{s_i-1}{r_i}} \right) = \widetilde{x}_{s_i}^{(i)}, \text{ say,}$$

for  $i \in \{1, \dots, m\}$ . By plugging in  $\widetilde{x}_{s_i}^{(i)}, i = 1, \dots, m$ , we arrive at

$$\begin{aligned} & \left( 1 - \frac{\ln \left( 1 - F_\theta \left( x_{r_i}^{(i)} \right) \right)}{\ln \left( 1 - F_\theta \left( \tilde{x}_{s_i}^{(i)} \right) \right)} \right)^{s_i - r_i - 1} \left( \frac{\ln \left( 1 - F_\theta \left( x_{r_i}^{(i)} \right) \right)}{\ln \left( 1 - F_\theta \left( \tilde{x}_{s_i}^{(i)} \right) \right)} \right)^{r_i} \\ & = \left( 1 - \frac{r_i}{s_i - 1} \right)^{s_i - r_i - 1} \left( \frac{r_i}{s_i - 1} \right)^{r_i}, \end{aligned} \tag{3}$$

which does not depend on the unknown parameter  $\theta$ .

If  $s_i = r_i + 1$  for an  $i \in \{1, \dots, m\}$ , the above function is decreasing in  $x_{s_i}^{(i)}$  and therefore maximized in  $x_{s_i}^{(i)} = \tilde{x}_{s_i}^{(i)} = x_{r_i}^{(i)}$ .

Thus, the remaining function  $\psi$  still has to be maximized with respect to  $\theta$ . □

If, for an  $i \in \{1, \dots, m\}$ ,  $s_i = r_i + 1$ , i.e., the prediction of the next record is of interest, then the marginal MOLP equals  $R_{r_i}^{(i)}$ , which denotes the last observed record; hence, the predictor is not useful. This is pointed out in Volovskiy and Kamps (2020a), where the one-sample case  $m = 1$  is examined, which is included, here.

In order to extend the MPSP from the one-sample case as in Volovskiy and Kamps (2020b) to a simultaneous MPSP in the case of  $m$  independent samples, we consider the function

$$\begin{aligned} & \mathcal{P}_{s_1, \dots, s_m} \left( x_1^{(1)}, \dots, x_{s_1}^{(1)}, \dots, x_1^{(m)}, \dots, x_{s_m}^{(m)}, \theta \right) \\ & = \prod_{i=1}^m \left( \prod_{j=1}^{s_i} \left( \frac{H_\theta \left( x_j^{(i)} \right)}{H_\theta \left( x_{s_i}^{(i)} \right)} - \frac{H_\theta \left( x_{j-1}^{(i)} \right)}{H_\theta \left( x_{s_i}^{(i)} \right)} \right) \mathbb{I}_{(\alpha(F_\theta), \omega(F_\theta))} \left( x_{s_i}^{(i)} \right) \right), \end{aligned} \tag{4}$$

for  $r_i, s_i \in \mathbb{N}$  with  $r_i \geq 2$  and  $s_i > r_i, i = 1, \dots, m, x_0^{(i)} = \alpha(F_\theta)$  for all  $i = 1, \dots, m$ , where maximization is with respect to  $x_{r_1+1}^{(1)}, \dots, x_{s_1}^{(1)}, \dots, x_{r_m+1}^{(m)}, \dots, x_{s_m}^{(m)}$  and  $\theta$ .

As in Theorem 1 for the simultaneous MOLP, a representation of the simultaneous MPSP can be stated subject to the existence of a maximum of a certain function regarding  $\theta$ .

**Theorem 2** *If, for  $r_i, s_i \in \mathbb{N}$  with  $r_i \geq 2$  and  $s_i \geq r_i + 1, i = 1, \dots, m, m \in \mathbb{N}, a$  Borel measurable function  $\widehat{\theta}: \mathbb{R}^{\sum_{i=1}^m r_i} \rightarrow \Theta$  exists with*

$$\mathcal{P}_{r_1, \dots, r_m} \left( x_\star^{(1)}, \dots, x_\star^{(m)}, \widehat{\theta} \left( x_\star^{(1)}, \dots, x_\star^{(m)} \right) \right) = \max_{\theta \in \Theta} \mathcal{P}_{r_1, \dots, r_m} \left( x_\star^{(1)}, \dots, x_\star^{(m)}, \theta \right)$$

for any  $x_\star^{(i)} \in \mathbb{R}_{<}^{r_i}, i = 1, \dots, m$ , then the MPSP of  $R_{s_i}^{(i)}$  has the form

$$\pi_{MPSP}^{(i)} \left( R_\star^{(1)}, \dots, R_\star^{(m)} \right) = F_{\widehat{\theta} \left( x_\star^{(1)}, \dots, x_\star^{(m)} \right)}^{-1} \left( 1 - \left( 1 - F_{\widehat{\theta} \left( x_\star^{(1)}, \dots, x_\star^{(m)} \right)} \left( R_{r_i}^{(i)} \right) \right)^{\frac{s_i}{r_i}} \right)$$

for all  $i = 1, \dots, m$  simultaneously.

**Proof** By inspecting the proof of Theorem 1 in Volovskiy and Kamps (2020b) for the case  $m = 1$ , and with a suitable constant  $c$  that depends only on  $r_1, \dots, r_m, s_1, \dots, s_m$ , we obtain:

$$\begin{aligned} & \max_{(y^{(1)}, \dots, y^{(m)}, \theta) \in (\mathbb{R}_{<}^{s_1-1} \times \dots \times \mathbb{R}_{<}^{s_m-r_m} \times \Theta)} : (x_{\star}^{(i)}, y^{(i)}) \in \mathbb{R}_{<}^{s_i} \mathcal{P}_{s_1, \dots, s_m} (x_{\star}^{(1)}, y^{(1)}, \dots, x_{\star}^{(m)}, y^{(m)}, \theta) \\ &= \max_{\theta \in \Theta} \prod_{i=1}^m \max_{x_{s_i}^{(i)} \in (x_{r_i}^{(i)}, \infty)} \max_{(x_{r_i+1}^{(i)}, \dots, x_{s_i-1}^{(i)}) \in (x_{r_i}^{(i)}, x_{s_i}^{(i)})_{<}} \left( \prod_{j=1}^{s_i} \left( \frac{H_{\theta}(x_j^{(i)})}{H_{\theta}(x_{s_i}^{(i)})} - \frac{H_{\theta}(x_{j-1}^{(i)})}{H_{\theta}(x_{s_i}^{(i)})} \right) \right) \\ &= \max_{\theta \in \Theta} \prod_{i=1}^m \max_{x_{s_i}^{(i)} \in (x_{r_i}^{(i)}, \infty)} \left( \prod_{j=1}^{r_i} \left( \frac{H_{\theta}(x_j^{(i)})}{H_{\theta}(x_{s_i}^{(i)})} - \frac{H_{\theta}(x_{j-1}^{(i)})}{H_{\theta}(x_{s_i}^{(i)})} \right) \right) \left( \frac{1 - H_{\theta}(x_{r_i}^{(i)}) / H_{\theta}(x_{s_i}^{(i)})}{s_i - r_i} \right)^{s_i - r_i} \\ &= c \max_{\theta \in \Theta} \mathcal{P}_{r_1, \dots, r_m} (x_{\star}^{(1)}, \dots, x_{\star}^{(m)}, \theta) \prod_{i=1}^m \max_{x_{s_i}^{(i)} \in (x_{r_i}^{(i)}, \infty)} \left( \frac{H_{\theta}(x_{r_i}^{(i)})}{H_{\theta}(x_{s_i}^{(i)})} \right)^{r_i} \left( 1 - \frac{H_{\theta}(x_{r_i}^{(i)})}{H_{\theta}(x_{s_i}^{(i)})} \right)^{s_i - r_i}. \end{aligned}$$

The factors of the latter product take the shape of a beta density functions with mode

$$\frac{H_{\theta}(x_{r_i}^{(i)})}{H_{\theta}(x_{s_i}^{(i)})} = \frac{r_i}{s_i}, \quad \text{for all } i = 1, \dots, m.$$

Hence, because of the product structure the maximum of  $\mathcal{P}_{s_1, \dots, s_m}$  is attained at

$$x_{s_i}^{(i)} = F_{\theta}^{-1} \left( 1 - \left( 1 - F_{\theta}(x_{r_i}^{(i)}) \right)^{\frac{s_i}{r_i}} \right)$$

for all  $i = 1, \dots, m$  and it remains to maximize  $\mathcal{P}_{r_1, \dots, r_m}$  regarding  $\theta \in \Theta$ . □

As in the one-sample setting in Volovskiy and Kamps (2020b) the MPSP allows a reasonable prediction of the very next record in every sequence and the case  $m = 1$  is included in Theorem 2.

Point predictors turn out to be helpful in the development of prediction intervals for future record values. Prediction intervals for records have been studied in the case of a single sample for various underlying distributions (see, e.g., Awad and Raqab (2000), Raqab (2007)). In Empacher et al (2024), the MPSP for a future record value based on a Pareto distribution (see Theorem 2 for  $m = 1$ , Volovskiy and Kamps (2020b)) has been successfully utilized to construct prediction intervals, which perform well in terms of expected length and coverage. By means of simultaneous predictors in  $m$  independent samples,  $m$ -dimensional rectangular prediction regions are near at hand.

### 3 Record values from particular distributions

In Theorem 1 and Theorem 2 the general forms of the MOLP and the MPSP are presented, which, however, are usually not explicitly available. For record values with

underlying exponential or Pareto distributions, the predictors MOLP, MPSP as well as MLP can be derived in closed form as shown in the following sections.

### 3.1 Exponential distribution

Let the two-parameter exponential distribution with location parameter  $\mu \in \mathbb{R}$  and scale parameter  $\sigma \in (0, \infty)$  be given by the cdf

$$F_{\mu, \sigma}(x) = 1 - \exp\left(-\frac{x - \mu}{\sigma}\right), \quad x \in (\mu, \infty). \quad (5)$$

For point prediction of future record values based on an exponential distribution in the single sample case, we refer to Basak and Balakrishnan (2003), Raqab et al (2007), Volovskiy and Kamps (2020a, b), Almuzaini and Raqab (2022), Navarro (2022).

According to Theorem 1 we have

$$\Psi\left(\mu, \sigma \mid x_{\star}^{(1)}, \dots, x_{\star}^{(m)}\right) = \prod_{i=1}^m \frac{1}{\left(x_{r_i}^{(i)} - \mu\right)^{r_i}} \mathbb{I}_{\left(0, \min\{x_1^{(1)}, \dots, x_1^{(m)}\}\right)}(\mu).$$

Since  $\Psi$  does not depend on  $\sigma$  and it is increasing in  $\mu$ ,  $\Psi$  is maximized in  $\widehat{\mu}_{MOL} = \min\{x_1^{(1)}, \dots, x_1^{(m)}\}$ . Hence, by Theorem 1, the MOLP is given by

$$F_{\mu, \sigma}^{-1}\left(1 - \left(1 - F_{\mu, \sigma}\left(R_{r_i}^{(i)}\right)\right)^{\frac{s_i - 1}{r_i}}\right) = R_{r_i}^{(i)} + \left(R_{r_i}^{(i)} - \mu\right) \frac{s_i - r_i - 1}{r_i}.$$

This yields the following result for the MOLP, which, in the one-sample case (i.e.,  $m = 1$ ), is shown in Volovskiy and Kamps (2020a).

**Corollary 1 (MOLP)** Let  $R_1^{(i)}, \dots, R_{s_i}^{(i)}$  be the first  $s_i \in \mathbb{N}$ ,  $i = 1, \dots, m$ , upper record values of  $m$  independent samples of random variables with cdf  $F_{\mu, \sigma}$  as in (5).

For  $r_i, s_i \in \mathbb{N}$  with  $2 \leq r_i < s_i$  the unique MOLP of  $R_{s_i}^{(i)}$  based on  $R_{\star}^{(1)}, \dots, R_{\star}^{(m)}$  is given by

$$\pi_{MOLP}^{(i)} = R_{r_i}^{(i)} + \left(R_{r_i}^{(i)} - \mu\right) \frac{s_i - r_i - 1}{r_i} \quad \text{for all } i = 1, \dots, m, \text{ simultaneously,} \quad (6)$$

if  $\mu$  is known, and by (6) with  $\mu$  replaced by  $\widehat{\mu}_{MOL} = \min\{R_1^{(1)}, \dots, R_1^{(m)}\}$ , if  $\mu$  is unknown.

For the construction of the MPSP we consider the function in (4) for the two-parameter exponential distribution as in (5):

$$\mathcal{P}_{r_1, \dots, r_m}(\mu, \sigma \mid x_\star^{(1)}, \dots, x_\star^{(m)}) = \prod_{i=1}^m \left( \frac{x_1^{(i)} - \mu}{(x_{r_i}^{(i)} - \mu)^{r_i}} \prod_{j=2}^{r_i} (x_j^{(i)} - x_{j-1}^{(i)}) \right),$$

which does not depend on the unknown parameter  $\sigma$ . A global maximum regarding  $\mu$  exists since

$$\begin{aligned} \lim_{\mu \rightarrow -\infty} \ln \left( \mathcal{P}_{r_1, \dots, r_m}(\mu, \sigma \mid x_\star^{(1)}, \dots, x_\star^{(m)}) \right) &= -\infty, \\ \lim_{\mu \rightarrow \min\{x_1^{(1)}, \dots, x_1^{(m)}\}} \ln \left( \mathcal{P}_{r_1, \dots, r_m}(\mu, \sigma \mid x_\star^{(1)}, \dots, x_\star^{(m)}) \right) &= -\infty, \end{aligned}$$

but has to be calculated numerically. For  $\mu$  known and according to Theorem 2, the MPSP takes the form

$$F_{\mu, \sigma}^{-1} \left( 1 - \left( 1 - F_{\mu, \sigma} \left( R_{r_i}^{(i)} \right) \right)^{\frac{s_i}{r_i}} \right) = R_{r_i}^{(i)} + \left( R_{r_i}^{(i)} - \mu \right) \frac{s_i - r_i}{r_i}.$$

Thus, we find the following corollary.

**Corollary 2 (MPSP)** *Let  $R_1^{(i)}, \dots, R_{s_i}^{(i)}$  be the first  $s_i \in \mathbb{N}, i = 1, \dots, m$ , upper record values of  $m$  independent samples of random variables with cdf  $F_{\mu, \sigma}$  as in (5).*

*For  $r_i, s_i \in \mathbb{N}$  with  $2 \leq r_i < s_i$  the unique MPSP of  $R_{s_i}^{(i)}$  based on  $R_\star^{(1)}, \dots, R_\star^{(m)}$  is given by*

$$\pi_{MPSP}^{(i)} = R_{r_i}^{(i)} + \left( R_{r_i}^{(i)} - \mu \right) \frac{s_i - r_i}{r_i} \text{ for all } i = 1, \dots, m, \text{ simultaneously,} \quad (7)$$

if  $\mu$  is known, and by (7) with  $\mu$  replaced by

$$\hat{\mu}_{MPS} = \operatorname{argmax}_{\mu \in (0, \min\{x_1^{(1)}, \dots, x_1^{(m)}\})} \mathcal{P}_{r_1, \dots, r_m}(\mu, \sigma \mid x_\star^{(1)}, \dots, x_\star^{(m)}),$$

if  $\mu$  is unknown.

The one-sample MPSP as in Volovskiy and Kamps (2020b) is included by choosing  $m = 1$ .

Moreover, the simultaneous MLP of future sample record values can be stated explicitly when exponential distributions are assumed. In the single sample case, MLPs for known and unknown location parameter are shown in Basak and Balakrishnan (2003).

**Lemma 3 (MLP)** Let  $R_1^{(i)}, \dots, R_{s_i}^{(i)}$  be the first  $s_i \in \mathbb{N}, i = 1, \dots, m$ , upper record values of  $m$  independent samples of random variables with cdf  $F_{\mu, \sigma}$  as in (5). For  $r_i, s_i \in \mathbb{N}$  with  $2 \leq r_i < s_i$ , the unique MLP of  $R_{s_i}^{(i)}$  based on  $R_{\star}^{(1)}, \dots, R_{\star}^{(m)}$  is given by

$$\pi_{MLP}^{(i)} = R_{r_i}^{(i)} + \left( \sum_{i=1}^m R_{r_i}^{(i)} - m\mu \right) \frac{s_i - r_i - 1}{\sum_{i=1}^m r_i + m} \text{ for all } i = 1, \dots, m, \text{ simultaneously,} \tag{8}$$

if  $\mu$  is known, and by (8) with  $\mu$  replaced by  $\widehat{\mu}_{PML} = \min \{ R_1^{(1)}, \dots, R_1^{(m)} \}$ , if  $\mu$  is unknown.

**Proof** According to Equation (2) the predictive likelihood function is given by

$$\begin{aligned} PL(x_{\star}^{(1)}, x_{s_1}^{(1)}, \dots, x_{\star}^{(m)}, x_{s_m}^{(m)}, \mu, \sigma) &= \prod_{i=1}^m \left( \frac{\sigma^{-s_i}}{(s_i - r_i - 1)!} \exp\left(-\frac{x_{s_i}^{(i)} - \mu}{\sigma}\right) (x_{s_i}^{(i)} - x_{r_i}^{(i)})^{s_i - r_i - 1} \right) \\ &\propto \sigma^{-\sum_{i=1}^m s_i} \prod_{i=1}^m h_i(x_{s_i}^{(i)}, \sigma) \end{aligned} \tag{9}$$

with functions  $h_i$  given by

$$h_i(x_{s_i}^{(i)}, \mu, \sigma) = \exp\left(-\frac{x_{s_i}^{(i)} - \mu}{\sigma}\right) (x_{s_i}^{(i)} - x_{r_i}^{(i)})^{s_i - r_i - 1}, \quad i = 1, \dots, m.$$

Every  $h_i$  and therefore (9) is increasing in  $\mu$ . Hence, (9) is maximized with respect to  $\mu$  in  $\widehat{\mu}_{PML} = \min \{ x_1^{(1)}, \dots, x_m^{(1)} \}$ , which does not depend on  $\sigma$  or  $x_{s_i}^{(i)}$ .

Because of the product structure, each  $h_i$  can be maximized regarding  $x_{s_i}^{(i)}$  for fixed  $\sigma \in (0, \infty)$  separately.

If  $s_i = r_i + 1$ , then  $h_i$  is decreasing in  $x_{s_i}^{(i)}$  and therefore maximized in  $x_{r_i}^{(i)} = \widetilde{x}_{s_i}^{(i)}$ , say, with

$$h_i(x_{r_i}^{(i)}, \sigma) = \exp\left(-\frac{x_{r_i}^{(i)} - \mu}{\sigma}\right) = \exp\left(-\frac{x_{r_i}^{(i)} - \mu}{\sigma}\right) \sigma^{s_i - r_i - 1}.$$

If  $s_i > r_i + 1$ , then, for fixed  $\sigma \in (0, \infty)$ ,  $h_i$  is seen to attain its global maximum regarding  $x_{s_i} \in (x_{r_i}^{(i)}, \infty)$  at

$$x_{r_i}^{(i)} + \sigma(s_i - r_i - 1) = \widetilde{x}_{s_i}^{(i)}, \text{ say, with}$$

$$h_i \left( x_{r_i}^{(i)} + \sigma (s_i - r_i - 1), \sigma \right) \propto \exp \left( -\frac{x_{r_i}^{(i)} - \mu}{\sigma} \right) \sigma^{s_i - r_i - 1}.$$

Hence we find

$$PL \left( x_{\star}^{(1)}, \tilde{x}_{s_1}^{(1)}, \dots, x_{\star}^{(m)}, \tilde{x}_{s_m}^{(m)}, \sigma \right) \propto \sigma^{-\sum_{i=1}^m r_i - m} \exp \left( -\frac{\sum_{i=1}^m x_{r_i}^{(i)} - m\mu}{\sigma} \right) = g(\sigma), \text{ say.}$$

Moreover,  $g$  attains its global maximum with respect to  $\sigma \in (0, \infty)$  at

$$\sigma = \frac{\sum_{i=1}^m x_{r_i}^{(i)} - m\mu}{\sum_{i=1}^m r_i + m}.$$

Inspecting the derivation, we find uniqueness of the maximum of the predictive likelihood function and the form of the MLP to be

$$x_{s_i}^{(i)} = x_{r_i}^{(i)} + \left( \sum_{i=1}^m x_{r_i}^{(i)} - m\mu \right) \frac{s_i - r_i - 1}{\sum_{i=1}^m r_i + m}, \quad i = 1, \dots, m.$$

□

If, for some  $i \in \{1, \dots, m\}$ ,  $s_i = r_i + 1$ , then MLP yields a trivial predictor as in the one-sample case, i.e.,  $\pi_{MLP}^{(i)} = R_{r_i}^{(i)}$  namely the last observed record. Thus, there is no reasonable MLP, if the very next record is predicted.

### 3.2 Pareto distribution

Let the Pareto distribution with scale parameter  $\lambda \in (0, \infty)$  and shape parameter  $\beta \in (0, \infty)$  be given by the cdf

$$F_{\lambda, \beta}(x) = 1 - \left( \frac{\lambda}{x} \right)^\beta, \quad x \in [\lambda, \infty). \tag{10}$$

Point prediction of future Pareto record values in the single sample case has been studied in, e.g., Madi and Raqab (2004), Ahmadi and Doostparast (2006), Raqab et al (2007), and Volovskiy and Kamps (2020b, 2023).

In order to obtain the MOLP according to Theorem 1, we maximize the function  $\Psi$  given by

$$\Psi \left( \lambda, \beta \mid x_{\star}^{(1)}, \dots, x_{\star}^{(m)} \right) = \prod_{i=1}^m \frac{1}{\left( \ln \left( x_{r_i}^{(i)} \right) - \ln(\lambda) \right)^{r_i}} \prod_{j=1}^{r_i} \frac{1}{x_j^{(i)}} \mathbb{I}_{\left( 0, \min \{ x_1^{(1)}, \dots, x_1^{(m)} \} \right)}(\lambda)$$

with respect to  $\lambda$  and  $\beta$ .

However,  $\Psi$  does not depend on  $\beta$  and is increasing in  $\lambda$ . Therefore, for  $\lambda$  unknown, it has to be chosen as  $\widehat{\lambda}_{MOL} = \min \{x_1^{(1)}, \dots, x_1^{(m)}\}$ . The form of the MOLP is given by Theorem 1 as

$$F_{\lambda, \beta}^{-1} \left( 1 - \left( 1 - F_{\lambda, \beta} \left( R_{r_i}^{(i)} \right) \right)^{s_i - 1} r_i \right) = \lambda \left( \frac{R_{r_i}^{(i)}}{\lambda} \right)^{\frac{s_i - 1}{r_i}},$$

which yields the following result.

**Corollary 3 (MOLP)** Let  $R_1^{(i)}, \dots, R_{s_i}^{(i)}$  be the first  $s_i \in \mathbb{N}, i = 1, \dots, m$ , upper record values of  $m$  independent samples of random variables with cdf  $F_{\lambda, \beta}$  as in (10). For  $r_i, s_i \in \mathbb{N}$  with  $2 \leq r_i < s_i$  the unique MOLP of  $R_{s_i}^{(i)}$  based on  $R_{r_i}^{(1)}, \dots, R_{r_i}^{(m)}$  is given by

$$\pi_{MOLP}^{(i)} = \lambda \left( \frac{R_{r_i}^{(i)}}{\lambda} \right)^{\frac{s_i - 1}{r_i}} \quad \text{for all } i = 1, \dots, m, \text{ simultaneously,} \quad (11)$$

if  $\lambda$  is known, and by (11) with  $\lambda$  replaced by  $\widehat{\lambda}_{MOL} = \min \{R_1^{(1)}, \dots, R_1^{(m)}\}$ , if  $\lambda$  is unknown.

The single sample case  $m = 1$  is contained in Corollary 3 and is shown in Volovskiy and Kamps (2023) for an unknown parameter  $\lambda$ . It has to be pointed out that the marginal MOLP in the situation  $s_i = r_i + 1$  for some  $i \in \{1, \dots, m\}$  becomes trivial in the sense that the MOLP reduces to  $R_{r_i}^{(i)}$ , i.e., the last observed record value.

Next, the MPSP in the Pareto setting is derived. For the single sample case we refer to Volovskiy and Kamps (2020b). According to Theorem 2,

$$\begin{aligned} \mathcal{P}_{r_1, \dots, r_m}(\lambda, \beta \mid x_{\star}^{(1)}, \dots, x_{\star}^{(m)}) &= \prod_{i=1}^m \frac{\ln(x_1^{(i)}) - \ln(\lambda)}{(\ln(x_{r_i}^{(i)}) - \ln(\lambda))^{r_i}} \\ &\quad \bullet \prod_{j=2}^{r_i} (\ln(x_j^{(i)}) - \ln(x_{j-1}^{(i)})) \end{aligned}$$

with  $x_0^{(i)} = \lambda, i = 1, \dots, m$  has to be maximized regarding the parameters of the Pareto distribution. This function does not depend on the shape parameter  $\beta$ . Since  $\mathcal{P}_{r_1, \dots, r_m}$  is a non-negative function and

$$\begin{aligned} \lim_{\lambda \rightarrow 0} \mathcal{P}_{r_1, \dots, r_m}(\lambda, \beta \mid x_{\star}^{(1)}, \dots, x_{\star}^{(m)}) &= 0, \\ \lim_{\lambda \rightarrow \min \{x_1^{(1)}, \dots, x_1^{(m)}\}} \mathcal{P}_{r_1, \dots, r_m}(\lambda, \beta \mid x_{\star}^{(1)}, \dots, x_{\star}^{(m)}) &= 0, \end{aligned}$$

it attains a global maximum with respect to  $\lambda$ , which can be derived numerically, and which may not be unique. For  $\lambda$  known and because of Theorem 2, the MPSP takes the form

$$F_{\lambda,\beta}^{-1} \left( 1 - \left( 1 - F_{\lambda,\beta} \left( R_{r_i}^{(i)} \right) \right)^{\frac{s_i}{r_i}} \right) = \lambda \left( \frac{R_{r_i}^{(i)}}{\lambda} \right)^{\frac{s_i}{r_i}},$$

which leads to the following corollary.

**Corollary 4 (MPSP)** *Let  $R_1^{(i)}, \dots, R_{s_i}^{(i)}$  be the first  $s_i \in \mathbb{N}, i = 1, \dots, m$ , upper record values of  $m$  independent samples of random variables with cdf  $F_{\lambda,\beta}$  as in (10). For  $r_i, s_i \in \mathbb{N}$  with  $2 \leq r_i < s_i$  the unique MPSP of  $R_{s_i}^{(i)}$  based on  $R_{r_i}^{(1)}, \dots, R_{r_i}^{(m)}$  is given by*

$$\pi_{MPSP}^{(i)} = \lambda \left( \frac{R_{r_i}^{(i)}}{\lambda} \right)^{\frac{s_i}{r_i}} \quad \text{for all } i = 1, \dots, m, \text{ simultaneously,} \tag{12}$$

if  $\lambda$  is known, and by (12) with  $\lambda$  replaced by

$$\widehat{\lambda}_{MPS} = \operatorname{argmax}_{\lambda \in (0, \min\{x_1^{(1)}, \dots, x_1^{(m)}\})} \mathcal{P}_{r_1, \dots, r_m} \left( \lambda, \beta \mid x_{r_1}^{(1)}, \dots, x_{r_m}^{(m)} \right),$$

if  $\lambda$  is unknown.

In the single sample situation, the MLP for a future Pareto record value is shown in Raqab et al (2007), and Volovskiy and Kamps (2023) for an unknown parameter  $\lambda$ . This result can be extended to the multi sample setting. The following lemma states the simultaneous MLP for both, known and unknown scale parameter.

**Lemma 4 (MLP)** *Let  $R_1^{(i)}, \dots, R_{s_i}^{(i)}$  be the first  $s_i \in \mathbb{N}, i = 1, \dots, m$ , upper record values of  $m$  independent samples of random variables with cdf  $F_{\lambda,\beta}$  as in (10).*

*For  $r_i, s_i \in \mathbb{N}$  with  $2 \leq r_i < s_i$  for all  $i = 1, \dots, m$ , the unique MLP of  $R_{s_i}^{(i)}$  based on  $R_{r_i}^{(1)}, \dots, R_{r_i}^{(m)}$  is given by*

$$\pi_{MLP}^{(i)} = R_{r_i}^{(i)} \exp \left( \frac{s_i - r_i - 1}{\widehat{\beta}_{PML} + 1} \right) \tag{13}$$

with

$$\widehat{\beta}_{PML} = \frac{1}{2} \left( \frac{\sum_{i=1}^m r_i + m}{\sum_{i=1}^m \ln \left( \frac{R_{r_i}^{(i)}}{\lambda} \right)} - 1 + \sqrt{\left( \frac{\sum_{i=1}^m r_i + m}{\sum_{i=1}^m \ln \left( \frac{R_{r_i}^{(i)}}{\lambda} \right)} - 1 \right)^2 + \frac{4 \sum_{i=1}^m s_i}{\sum_{i=1}^m \ln \left( \frac{R_{r_i}^{(i)}}{\lambda} \right)}} \right),$$

if  $\lambda$  is known, and by (13) with  $\lambda$  in the representation of  $\widehat{\beta}_{PML}$  replaced by  $\widehat{\lambda}_{PML} = \min \{R_1^{(1)}, \dots, R_1^{(m)}\}$ , if  $\lambda$  is unknown.

**Proof** According to Equation (2) the predictive likelihood function  $PL(x_\star^{(1)}, x_{s_1}^{(1)}, \dots, x_\star^{(m)}, x_{s_m}^{(m)}, \beta)$  is proportional to

$$\lambda^{m\beta} \beta^{\sum_{i=1}^m s_i} \prod_{i=1}^m h_i(x_{s_i}^{(i)}, \beta)$$

with functions  $h_i$  given by

$$h_i(x_{s_i}^{(i)}, \beta) = (x_{s_i}^{(i)})^{-(\beta+1)} \left( \ln(x_{s_i}^{(i)}) - \ln(x_{r_i}^{(i)}) \right)^{s_i-r_i-1}, \quad i = 1, \dots, m.$$

The  $h_i$  and therefore the predictive likelihood function are increasing in  $\lambda$  and maximized in  $\lambda = \min \{x_1^{(1)}, \dots, x_1^{(m)}\}$ .

Because of the product structure, each  $h_i$  can be maximized regarding  $x_{s_i}^{(i)}$  for fixed  $\beta \in (0, \infty)$  separately.

If  $s_i = r_i + 1$ , then  $h_i$  is decreasing in  $x_{s_i}^{(i)}$  and therefore maximized in  $x_{r_i}^{(i)} = \widetilde{x}_{s_i}^{(i)}$ , say, with

$$h_i(\widetilde{x}_{s_i}^{(i)}, \beta) = (x_{r_i}^{(i)})^{-(\beta+1)}.$$

If  $s_i > r_i + 1$ , then, for fixed  $\beta \in (0, \infty)$ ,  $h_i$  is seen to attain its global maximum regarding  $x_{s_i}^{(i)} \in (x_{r_i}^{(i)}, \infty)$  at

$$x_{r_i}^{(i)} \exp\left(\frac{s_i - r_i - 1}{\beta + 1}\right) = \widetilde{x}_{s_i}^{(i)}, \text{ say, with}$$

$$h_i(\widetilde{x}_{s_i}^{(i)}, \beta) \propto \frac{1}{(x_{r_i}^{(i)})^{\beta+1}} \frac{1}{(\beta + 1)^{s_i-r_i-1}}.$$

Thus, we obtain

$$\begin{aligned} & PL(x_\star^{(1)}, \widetilde{x}_{s_1}^{(1)}, \dots, x_\star^{(m)}, \widetilde{x}_{s_m}^{(m)}, \beta) \\ & \propto \lambda^{m\beta} \beta^{\sum_{i=1}^m s_i} \left( \prod_{i=1}^m x_{r_i}^{(i)} \right)^{-(\beta+1)} (\beta + 1)^{-\sum_{i=1}^m (s_i-r_i-1)} = g(\beta), \text{ say,} \end{aligned}$$

and

$$\begin{aligned} \ln(g(\beta)) &= m\beta \ln(\lambda) + \ln(\beta) \sum_{i=1}^m s_i - (\beta + 1) \sum_{i=1}^m \ln(x_{r_i}^{(i)}) \\ &\quad - \ln(\beta + 1) \sum_{i=1}^m (s_i - r_i - 1) \\ &= \beta \sum_{i=1}^m \ln\left(\frac{\lambda}{x_{r_i}^{(i)}}\right) \\ &\quad + \ln\left(\frac{\beta}{\beta + 1}\right) \sum_{i=1}^m s_i + \ln(\beta + 1) \\ &\quad \sum_{i=1}^m (r_i + 1) - \sum_{i=1}^m \ln(x_{r_i}^{(i)}). \end{aligned}$$

Since

$$\lim_{\beta \rightarrow 0} \ln(g(\beta)) = -\infty, \quad \lim_{\beta \rightarrow \infty} \ln(g(\beta)) = -\infty$$

$g$  attains its global maximum regarding  $\beta \in (0, \infty)$  at  $\widehat{\beta}$  (as given in the statement of Lemma 4), which is uniquely derived by solving

$$\frac{\partial \ln(g(\beta))}{\partial \beta} = m \ln(\lambda) + \frac{\sum_{i=1}^m s_i}{\beta} - \sum_{i=1}^m \ln(x_{r_i}^{(i)}) - \frac{\sum_{i=1}^m (s_i - r_i - 1)}{\beta + 1} = 0$$

By plugging in this maximizer of  $g$  in the formula for  $x_{s_i}^{(i)}$ , the form of the MLP is given by

$$x_{s_i}^{(i)} = x_{r_i}^{(i)} \exp\left(\frac{s_i - r_i - 1}{\widehat{\beta} + 1}\right), \quad i = 1, \dots, m.$$

□

As in the previous section in case of exponential distributions, the marginal MLP becomes trivial in the situation  $s_i = r_i + 1$  for some  $i \in \{1, \dots, m\}$  in the sense that  $\pi_{MLP}^{(i)} = R_{r_i}^{(i)}$ ; i.e., the last observed record value turns out to be the MLP for the next record value to appear.

In the following application of simultaneously predicting future record values in independent samples, we restrict ourselves to using the MPSP in order to study the most important case  $s_i = r_i + 1$  for all  $i \in \{1, \dots, m\}$ . As aforementioned, MOLP and MLP lead to trivial predictions by just using the last observed record value in each sample as predictor of the respective next record value to appear.

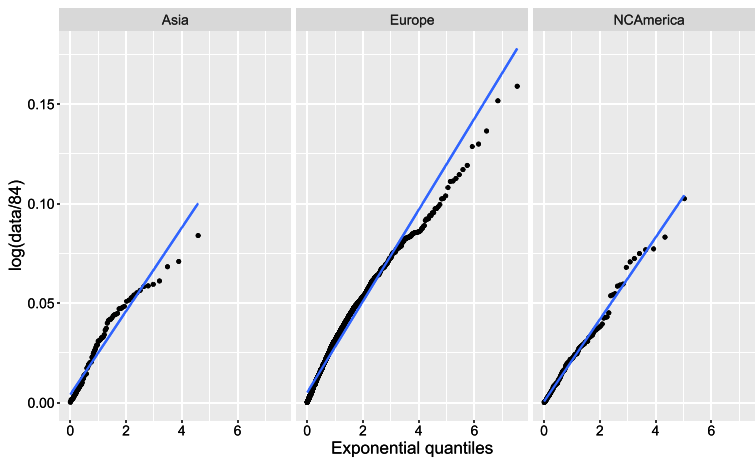


Fig. 1 Q-Q plot of the javelin throws above 84 m by continents

## 4 Application

As an application of records from several samples, continental athletic records in men's javelin throw are considered. The lengths of the javelin throws from April 1st, 1986 until 2023 can be found at <https://worldathletics.org/records/all-time-toplists/throws/javelin-throw/all/men/senior?regionType=world&page=1&bestResultsOnly=false&firstDay=1900-01-01&lastDay=2023-12-31&maxResultsByCountry=all&eventId=10229636&ageCategory=senior>. On April 1st, 1986 a regulation for a modified javelin came into effect, which led to shorter lengths compared to throws with the previous type of javelin.

It is possible to filter the data by area, such that for each continent records can be determined. In many situations, sports data above some threshold are quite well described by a Pareto distribution as mentioned in Empacher et al (2023). We will consider such an example with an appropriately chosen threshold  $\lambda$ , where only data above this threshold are taken into account. The javelin data is split into lengths before January 1st, 1987 and thereafter. The lengths of javelin throws in the first year after the javelin modification are used in order to evaluate, which lengths can be considered to indicate high-performance sport. Then, only lengths after 1986 and larger than 84 meters are considered further on to evaluate the fit to a Pareto distribution, and we restrict ourselves to lengths and records from athletes from Asia, Europe and North-and Central-America, here. The Pareto assumption is visualized by the quantile-quantile plots in Figure 1, which indicate a reasonable fit of the underlying data to a Pareto distribution. The slopes of the lines indicate the value of the shape parameter  $\beta$  which is similar for all three continents. Hence, we assume for illustration that the javelin throws from these continents can be described well by the same Pareto distribution.

When predicting records from several samples only records after January 1st, 1987 are taken into account, and we consider the value of  $\lambda$  unknown.

**Table 1** Prediction of continental records based on records between January 1st, 1987 and December 31st, 1995 in men’s javelin

<i>r</i>	Asia		Europe		North-/ Central-America	
	Athlete	Record	Athlete	Record	Athlete	Record
1	Mizoguchi	84.16	Železný	87.66	Petranoff	86.32
2	Yevsyukov	85.16	Bodén	89.10	Petranoff	86.46
3	Mizoguchi	85.22	Backley	89.58	Petranoff	89.12
4	Mizoguchi	87.60	Kinnunen	90.82	prediction	91.12 (90.71)
5	prediction	88.65 (89.15)	Backley	91.46		
6			Železný	95.54		
7			Železný	95.66		
8			prediction	97.53 (98.48)		

**Table 2** Prediction of continental records based on records between January 1st, 1987 and December 31th, 2023 in men’s javelin

<i>r</i>	Asia		Europe		North-/Central-America	
	Athlete	Record	Athlete	Record	Athlete	Record
1	Mizoguchi	84.16	Železný	87.66	Petranoff	86.32
2	Yevsyukov	85.16	Bodén	89.10	Petranoff	86.46
3	Mizoguchi	85.22	Backley	89.58	Petranoff	89.16
4	Mizoguchi	87.60	Kinnunen	90.82	Greer	90.71
5	Zhao	89.15	Backley	91.46	Greer	91.29
6	Cheng	91.36	Železný	95.54	Peters	93.07
7	prediction	92.75	Železný	95.66	prediction	94.78
8			Železný	98.48		
9			prediction	100.54		

In this application we distinguish two cases. First, we take a point of view as on December 31st, 1995 and consider records that have been observed until that point in time, only, in order to illustrate simultaneous prediction and to match predicted values and realized subsequent record values. The observed records and MPSP for each continent can be found in Table 1. The actual performed next record after December 31st, 1995 is given in parentheses. The numerical optimization of

$$P_{r_1, r_2, r_3}(\lambda, \beta) = \prod_{i=1}^3 \frac{\ln(x_1^{(i)}) - \ln(\lambda)}{(\ln(x_{r_i}^{(i)}) - \ln(\lambda))^{r_i}} \prod_{j=2}^{r_i} (\ln(x_j^{(i)}) - \ln(x_{j-1}^{(i)}))$$

needed in Corollary 4 yields  $\widehat{\lambda}_{MPS} = 83.52$  as an estimate of the unknown parameter  $\lambda$ .

The MPSP of the Asian and North- and Central-American record are quite close to the actual performed record values. Although there are more observed record values in Europe, the prediction is about a meter apart from the actual performed European record of Jan Železný, which is the current world record as well.

More meaningful predictions are possible by determining one- or two-sided prediction intervals (cf. Empacher et al (2024) in the single-sample case).

Next, we take a more recent point of view and consider simultaneous prediction of record values as from December 31st, 2023. The prediction results for future continental records are shown in Table 2. Here, the prediction procedure yields an estimate of  $\hat{\lambda}_{MPS} = 83.46$ . The predicted next European record of 100.54 m is the only one that breaks the current world record of 98.48 m as well.

Simultaneous prediction of record values may also be of interest in other applications, where independent samples can be assumed and where a good fit to a particular distribution is seen to be valid in all samples under consideration.

**Funding** Open Access funding enabled and organized by Projekt DEAL. No funds, grants, or other support was received.

## Declarations

**Conflicts of Interest** On behalf of all authors, the corresponding author states that there is no conflict of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Ahmadi J, Balakrishnan N (2010) Prediction of order statistics and record values from two independent sequences. *Statistics* 44(4):417–430
- Ahmadi J, Doostparast M (2006) Bayesian estimation and prediction for some life distributions based on record values. *Stat Pap* 47(3):373–392
- Ahmadi J, MirMostafae SMTK (2009) Prediction intervals for future records and order statistics coming from two parameter exponential distribution. *Stat Probab Lett* 79(7):977–983
- Ahmadi J, MirMostafae SMTK, Balakrishnan N (2010) Nonparametric prediction intervals for future record intervals based on order statistics. *Stat Probab Lett* 80(21–22):1663–1672
- Ahsanullah M (1980) Linear prediction of record values for the two parameter exponential distribution. *Ann Inst Stat Math* 32:363–368
- Almuzaini HA, Raqab MZ (2022) Comparative study for assessing the Pitman's closeness of predictors based on exponential record data. *Commun Stat Simul Comput* 51(8):4427–4449
- Awad AM, Raqab MZ (2000) Prediction intervals for the future record values from exponential distribution: comparative study. *J Stat Comp Simul* 65(1–4):325–340
- Basak P, Balakrishnan N (2003) Maximum likelihood prediction of future record statistic. In: Lindqvist B, Doksum K (eds) *Mathematical and Statistical Methods in Reliability*. World Scientific Publishing, New York, chap 11, p 159–175

- Basiri E, Fernández AJ, Asgharzadeh A, Bagheri SF (2020) Best prediction regions for future exponential record intervals. *Statistics* 54(5):969–988
- Empacher C, Kamps U, Volovskiy G (2023) Statistical prediction of future sports records based on record values. *Stats* 6(1):131–147
- Empacher C, Kamps U, Schmiedt AB (2024) Prediction intervals for future Pareto record claims. *Eur Actuar J* 15:163–197
- Kaminsky KS, Rhodin LS (1985) Maximum likelihood prediction. *Ann Inst Stat Math* 37(3):507–517
- Madi MT, Raqab MZ (2004) Bayesian prediction of temperature records using the Pareto model. *Environmetrics* 15(7):701–710
- Navarro J (2022) Prediction of record values by using quantile regression curves and distortion functions. *Metrika* 85(6):675–706
- Raqab MZ (2007) Exponential distribution records: different methods of prediction. In: Ahsanullah M, Raqab MZ (eds) *Recent Developments in Ordered Random Variables*. Nova Science Publ, New York, NY, pp 239–251
- Raqab MZ (2009) Distribution-free prediction intervals for the future current record statistics. *Stat Pap* 50:429–439
- Raqab MZ, Balakrishnan N (2008) Prediction intervals for future records. *Stat Probab Lett* 78(13):1955–1963
- Raqab MZ, Ahmadi J, Doostparast M (2007) Statistical inference based on record data from Pareto model. *Statistics* 41(2):105–118
- Singh S, Tripathi YM, Wu SJ (2017) Bayesian estimation and prediction based on lognormal record values. *J Appl Stat* 44(5):916–940
- Volovskiy G, Kamps U (2020) Maximum observed likelihood prediction of future record values. *TEST* 29(4):1072–1097
- Volovskiy G, Kamps U (2020) Maximum product of spacings prediction of future record values. *Metrika* 83(7):853–868
- Volovskiy G, Kamps U (2023) Comparison of likelihood-based predictors of future Pareto and Lomax record values in terms of Pitman closeness. *Commun Stat Theory Methods* 52(6):1905–1922

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.