



Article

Statistical Prediction of Future Sports Records Based on Record Values

Christina Empacher , Udo Kamps *  and Grigoriy Volovskiy 

Institute of Statistics, RWTH Aachen University, D-52056 Aachen, Germany

* Correspondence: udo.kamps@rwth-aachen.de

Abstract: Point prediction of future record values based on sequences of previous lower or upper records is considered by means of the method of maximum product of spacings, where the underlying distribution is assumed to be a power function distribution and a Pareto distribution, respectively. Moreover, exact and approximate prediction intervals are discussed and compared with regard to their expected lengths and their percentages of coverage. The focus is on deriving explicit expressions in the point and interval prediction procedures. Predictions and forecasts are of interest, e.g., in sports analytics, which is gaining more and more attention in several sports disciplines. Previous works on forecasting athletic records have mainly been based on extreme value theory. The presented statistical prediction methods are exemplarily applied to data from various disciplines of athletics as well as to data from American football based on fantasy football points according to the points per reception scoring scheme. The results are discussed along with basic assumptions and the choice of underlying distributions.

Keywords: point and interval prediction; power function distribution; Pareto distribution; athletics; American football



Citation: Empacher, C.; Kamps, U.; Volovskiy, G. Statistical Prediction of Future Sports Records Based on Record Values. *Stats* **2023**, *6*, 131–147. <https://doi.org/10.3390/stats6010008>

Academic Editor: Wei Zhu

Received: 29 November 2022

Revised: 3 January 2023

Accepted: 6 January 2023

Published: 11 January 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The consideration, collection and importance of data and their analysis is gaining more and more attention in many different fields—as it does in several sports disciplines. For example, in US sports such as baseball, basketball and American football, data regarding many aspects of a game, team or player have been recorded for many years. This trend can be observed in other areas of sports such as athletics, as well. Thus, statistical methods and analysis have become a topic of growing interest, and an interesting part in sports analysis are predictions and forecasts.

Among many other authors and articles on statistical methodology in sports, we refer to some of them. Based on results for order statistics and record values, Noubary [1] proposed methods to forecast athletic records. Several authors applied extreme value theory. In Einmahl and Magnus [2], the authors examined and compared world records in athletics based on extreme value theory, and they forecast the ultimate record in a specific discipline. By this, they initiated a series of subsequent works in this area. Noubary [3] was concerned with athletic records and calculated the probabilities of future best performances based on tail estimation in extreme value theory. Einmahl and Smeets [4] estimated ultimate world records for the 100 m running of women and men. Henriques-Rodrigues et al. [5] focused on the semiparametric estimation of the extreme value index and of the right endpoint of the underlying distribution in case it was finite. The ultimate record in men's long jump was examined in Fraga Alves et al. [6]. Annual best times were used in Stephenson and Tawn [7] to fit a model based on extreme value distributions to come up with extrapolated values which served as predicted records. Adam and Tawn [8] considered annual records in athletics and swimming disciplines and fitted the parameters of the generalized extreme value distribution as underlying distribution.

A variety of other approaches for data from, e.g., baseball, football and athletics, can be found in the literature. As examples, we refer to Albert et al. [9] and Wunderlich and Memmert [10].

In contrast to forecasting procedures via extreme value theory in the sense of a long-term prediction by determining an extreme value distribution, the focus of our approach is on statistical prediction of the next record or records by means of previous record values in a particular sports discipline. We first repeat properties of upper and lower record values while focusing on lower records. We use the maximum product of spacings prediction procedure to derive a point predictor of the next record value based on previous ones, which are usually recorded. In the statistical model, we assume the lower/upper record values to be based on a sequence of independent and identically distributed random variables having a power function distribution/Pareto distribution due to our focus on explicit statistical prediction methods. Then, we derive some exact and approximate prediction intervals. We compare these intervals with regard to their expected lengths, and we analyze their percentages of coverage in a simulation study. We apply our results to real data from women's 100 m and various other disciplines of athletics as well as to American football data by considering the evaluation of players by assigning so-called fantasy football points to players' actions. Moreover, we discuss the distributional assumptions.

2. Prediction of Record Values

We first recall some basic properties of record values. For details, we refer to the monograph of Arnold et al. [11].

Let $(X_i)_{i \in \mathbb{N}}$ be a sequence of independent and identically distributed (iid) random variables with absolutely continuous cumulative distribution function (cdf) F and probability density function (pdf) f . Here, \mathbb{N} denotes the set of positive integers. The random variables

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

are called upper record times. The random variables $R_n = X_{T_u(n)}$, $n \in \mathbb{N}$, are then called upper record values. Analogously, the lower record times

$$T_l(1) = 1 \quad \text{and} \quad T_l(n+1) = \min\{j > T_l(n) : X_j < X_{T_l(n)}\}, \quad n \in \mathbb{N},$$

are used to define the sequence of lower record values $L_n = X_{T_l(n)}$, $n \in \mathbb{N}$. The joint densities for the first $r \in \mathbb{N}$ upper and lower record values $R_\star = (R_1, \dots, R_r)$ and $L_\star = (L_1, \dots, L_r)$ are given by

$$f_{R_\star}(x_1, \dots, x_r) = f(x_r) \prod_{i=1}^{r-1} \frac{f(x_i)}{1 - F(x_i)} \quad (1)$$

where $-\infty < x_1 < x_2 < \dots < x_r < \infty$ and

$$f_{L_\star}(x_1, \dots, x_r) = f(x_r) \prod_{i=1}^{r-1} \frac{f(x_i)}{F(x_i)} \quad (2)$$

where $\infty > x_1 > x_2 > \dots > x_r > -\infty$, respectively.

It is well known that, in distribution, the upper record values R_1, \dots, R_r as well as the lower record values L_1, \dots, L_r can be simultaneously expressed by independent and identically standard exponentially distributed random variables Z_1, \dots, Z_r via

$$\begin{aligned} (R_1, \dots, R_r) &\stackrel{d}{=} \left(F^{-1} \left(1 - \exp \left(- \sum_{i=1}^j Z_i \right) \right) \right)_{1 \leq j \leq r}, \\ (L_1, \dots, L_r) &\stackrel{d}{=} \left(F^{-1} \left(\exp \left(- \sum_{i=1}^j Z_i \right) \right) \right)_{1 \leq j \leq r}, \end{aligned} \quad (3)$$

where $\stackrel{d}{=}$ denotes equality in distribution (see Arnold et al. [11]).

Aiming at deriving explicit statistical procedures, we choose power function distributions in case of lower records and Pareto distributions in case of upper records as underlying distributions; both families of distributions turn out to have a good fit to various data sets in sports. For parameters $\lambda > 0$ and $\beta > 0$, the cdf and pdf of the power function distribution $Pow(\lambda, \beta)$ are given by

$$F_{\lambda, \beta}(x) = \left(\frac{x}{\lambda}\right)^{\beta} \quad \text{and} \quad f_{\lambda, \beta}(x) = \frac{\beta x^{\beta-1}}{\lambda^{\beta}}, \quad x \in (0, \lambda), \quad (4)$$

and, for parameters $\mu > 0$ and $\gamma > 0$, the cdf and pdf of the Pareto distribution $Par(\mu, \gamma)$ are given by

$$F_{\mu, \gamma}(x) = 1 - \left(\frac{\mu}{x}\right)^{\gamma} \quad \text{and} \quad f_{\mu, \gamma}(x) = \frac{\gamma \mu^{\gamma}}{x^{\gamma+1}}, \quad x \in (\mu, \infty). \quad (5)$$

In the sequel, we assume the parameters β and γ to be unknown and the (threshold) parameters λ and μ to be fixed.

There is a close relationship between power function and Pareto distributions, namely, if $Y \sim Par(\mu, \gamma)$ holds then $Y^{-1} \sim Pow(\lambda, \beta)$ follows, where $\lambda = \frac{1}{\mu}$ and $\beta = \gamma$. Because of this relationship, we get consistent results for running (looking for minimum times), throwing and jumping events in athletics (looking for maximum lengths, widths and heights) without transforming times in velocities, as, e.g., in Einmahl and Magnus [2].

2.1. Point Prediction

We consider point prediction of future record values based on previous records in a sequence of iid random variables. Among others, Kaminsky and Rhodin [12] presented the maximum likelihood predictor (MLP) and Volovskiy and Kamps [13] introduced the maximum observed likelihood prediction (MOLP). Based on a Pareto distribution, the MLP and MOLP for upper record values were shown in Volovskiy and Kamps [14]. However, both methods share the disadvantage that neither the next upper nor the next lower record value can be predicted reasonably, since the methods lead to the last record as predictor of the next one. Raqab et al. [15] derived the MLP and the so-called conditional median predictor. Moreover, there are some Bayesian results for estimation and prediction in the upper record Pareto setting in Ahmadi and Doostparast [16], Madi and Raqab [17] and Raqab et al. [15]. For further references, we refer to Volovskiy and Kamps [13] and Volovskiy and Kamps [14].

We applied the maximum product of spacings prediction (MPSP) procedure for lower record values along the lines of Volovskiy and Kamps [14] for upper record values. The idea underlying the MPSP method is to apply a transformation to the sequence $L_{\star} = (L_1, \dots, L_s)$ of observed records L_1, \dots, L_r , $r < s$ as well as the yet-to-be-observed record values up to L_s , such that the transformed random variables are distributed as order statistics $U_{1:s-1}, \dots, U_{s-1:s-1}$ from an iid sample of $s-1$ uniform random variables. Then, by considering the spacings of the observed data after applying the transformation, the MPSP method chooses the prediction value $\pi_{MPSP}^{(s)}$ for the record value L_s that renders the observed data as uniform as possible according to some measure of uniformity. For a related estimation method in parametric inference utilizing the probability integral transform, we refer to Cheng and Amin [18] and Ranneby [19].

In what follows, we assume that the observed record values are from a sequence of iid random variables with continuous cdf F_{θ} , which may depend on an unknown parameter $\theta \in \Theta$, where Θ denotes the set of parameters. Let G_{θ} denote the function $G_{\theta}(x) = -\ln(F_{\theta}(x))$, $x \in \mathbb{R}$. Then, Equation (3) reveals that

$$G_{\theta}(L_i) \stackrel{d}{=} \tilde{R}_i, \quad i \in \mathbb{N}, \quad (6)$$

where $(\tilde{R}_i)_{i=1}^\infty$ is the sequence of upper record values in an iid sequence of standard exponential random variables. Reasoning as in Volovskiy and Kamps [14], we conclude that

$$\frac{G_\theta(L_i)}{G_\theta(L_s)} \stackrel{d}{=} U_{i:(s-1)}, \quad i = 1, \dots, s-1. \quad (7)$$

As mentioned above, the prediction of the s th record is assumed to be based on observations of the record values L_1, \dots, L_r . Since the expected values of the order statistics $U_{1:s-1}, \dots, U_{s-1:s-1}$ are equidistantly arranged in the interval $(0, 1)$, Equation (7) motivates the following inferential procedure for the future realization of L_s : using the notational convention that, for an interval $I \subseteq \mathbb{R}$ and $n \in \mathbb{N}$, $I_{<}^n = \{(x_1, \dots, x_n) \in I^n \mid x_1 < x_2 < \dots < x_n\}$, for $n \in \mathbb{N}$, we define

$$\mathcal{Z}_n = \{(\theta, x_1, \dots, x_n) \in \Theta \times \mathbb{R}_{<}^n \mid (x_1, \dots, x_n) \in (\alpha(F_\theta), \omega(F_\theta))_{<}^n\}$$

and

$$\mathcal{P}_n(\theta, x_1, \dots, x_n) = \prod_{i=1}^n \left(\frac{G_\theta(x_i)}{G_\theta(x_n)} - \frac{G_\theta(x_{i-1})}{G_\theta(x_n)} \right), \quad (\theta, x_1, \dots, x_n) \in \mathcal{Z}_n,$$

and we call any function $\hat{x}^* = (\hat{x}_{r+1}, \dots, \hat{x}_s) : \mathbb{R}_{<}^r \rightarrow \mathbb{R}_{<}^{s-r}$ a maximum product of spacings predictor of L_s if, for a suitable function $\hat{\theta} : \mathbb{R}_{<}^r \rightarrow \Theta$, we have that \hat{x}^* and $\hat{\theta}$ maximize \mathcal{P}_s , i.e., for any $\theta \in \Theta$,

$$(\hat{\theta}(x_\star), x_\star, \hat{x}^*(x_\star)) \in \mathcal{Z}_s, \quad x_\star \in (\alpha(F_\theta), \omega(F_\theta))_{<}^r$$

and

$$\mathcal{P}_s(\hat{\theta}(x_\star), x_\star, \hat{x}^*(x_\star)) = \max_{\substack{\theta \in \Theta, x^* \in \mathbb{R}_{<}^{s-r}: \\ (\theta, x_\star, x^*) \in \mathcal{Z}_s}} \mathcal{P}_s(\theta, x_\star, x^*), \quad x_\star \in (\alpha(F_\theta), \omega(F_\theta))_{<}^r,$$

where $\alpha(F_\theta)$ and $\omega(F_\theta)$ denote the left and right endpoint of F_θ , respectively. Note that the maximization of \mathcal{P}_s with respect to θ is necessary since θ is considered to be unknown. However, given that in the present situation our focus is exclusively on predictive inference, the main object of interest is the point predictor given by the function $\pi_{MPSP}^{(s)} = \hat{x}^*$. By actually carrying out the above maximization, one obtains that the MPSP of L_s takes the form

$$\pi_{MPSP}^{(s)}(L_\star) = F_{\hat{\theta}(L_\star)}^{-1} \left(F_{\hat{\theta}(L_\star)}(L_r)^{\frac{s}{r}} \right), \quad s > r \quad (8)$$

where the function $\hat{\theta}$ is such that, for any $\theta \in \Theta$,

$$(\hat{\theta}(x_\star), x_\star) \in \mathcal{Z}_r, \quad x_\star \in (\alpha(F_\theta), \omega(F_\theta))_{<}^r$$

and

$$\mathcal{P}_r(\hat{\theta}(x_\star), x_\star) = \max_{\substack{\theta \in \Theta: \\ (\theta, x_\star) \in \mathcal{Z}_r}} \mathcal{P}_r(\theta, x_\star), \quad x_\star \in (\alpha(F_\theta), \omega(F_\theta))_{<}^r.$$

The proof of Equation (8) proceeds along the same lines as the proof of (Volovskiy and Kamps [14], Theorem 2.3). The maximum product of spacings prediction procedure is applied to the problem of sports records prediction in Section 3.

In the particular situation of an underlying power function distribution as in Equation (4), one easily sees that, for any sequence of observed lower record values $x_\star \in (0, \lambda)_{<}^r$, the objective function $\beta \mapsto \mathcal{P}_r(\beta, x_\star)$ is independent of the shape parameter β , since the ratio in Equation (7) is

$$\frac{G_\theta(L_i)}{G_\theta(L_s)} = \frac{\ln(L_i) - \ln(\lambda)}{\ln(L_s) - \ln(\lambda)}, \quad i = 1, \dots, s-1.$$

Consequently, according to Formula (8), the MPSP of L_s is given by

$$\pi_{MPSP}^{(s)}(L_*) = \lambda \left(\frac{L_r}{\lambda} \right)^{\frac{s}{r}}.$$

The MPSP is based on the ratios of G_θ evaluated for certain arguments. Hence, in the power function situation, the shape parameter β drops out.

Analogously, the MPSP of the s th upper record value in a Pareto distributed sequence of random variables is given by

$$\pi_{MPSP}^{(s)}(R_*) = \mu \left(\frac{R_r}{\mu} \right)^{\frac{s}{r}}.$$

It should be noted that the threshold parameters λ and μ are supposed to be known.

2.2. Interval Prediction

There are only a few articles in the literature dealing with prediction and prediction intervals for lower record values. In one of them, Wang et al. [20] discussed estimation and prediction in a lower record setting with a baseline cdf $F_{\lambda,\alpha}(x) = (D_\lambda(x))^\alpha$ for $x > 0$ with parameters $\lambda, \alpha > 0$ from the proportional reversed hazard family, where D_λ is a cdf itself depending only on the parameter λ .

Prediction intervals for upper record values have been studied in several articles. For example, Awad and Raqab [21] compared different intervals for records based on an exponential distribution. Asgharzadeh et al. [22] and Raqab et al. [15] introduced some intervals for upper record values from Pareto distributions. The intervals for lower record values proposed in this section are similar to those in the aforementioned articles and are applied to sports data in Section 3.

2.2.1. Lower Record Values from Power Function Distributions

According to Equation (2), the density of the first r lower record values is

$$f_{L_*}(x \mid \lambda, \beta) = \frac{\beta^r}{\lambda^\beta} x_r^{\beta-1} \prod_{i=1}^{r-1} \frac{1}{x_i}. \quad (9)$$

From the log-likelihood function

$$l(\beta \mid x, \lambda) = r \ln(\beta) - \beta \ln(\lambda) + (\beta - 1) \ln(x_r) - \sum_{i=1}^{r-1} \ln(x_i),$$

we obtain the maximum likelihood estimator

$$\hat{\beta}_{ML} = \frac{r}{\ln(\lambda) - \ln(L_r)}$$

of β , which is required in the sequel.

In order to derive some prediction intervals for future lower record values, we examine the statistic

$$T_1 = \frac{\ln(L_s) - \ln(L_r)}{\ln(L_r) - \ln(L_1)}.$$

By using Equation (3), one gets

$$T_1 \stackrel{d}{=} \frac{\ln(\lambda) - \frac{1}{\beta} \sum_{i=1}^s Z_i - \ln(\lambda) + \frac{1}{\beta} \sum_{i=1}^r Z_i}{\ln(\lambda) - \frac{1}{\beta} \sum_{i=1}^r Z_i - \ln(\lambda) + \frac{1}{\beta} Z_1} = \frac{\sum_{i=r+1}^s Z_i}{\sum_{i=2}^r Z_i}.$$

Since the Z_i are standard exponential random variables, the sums in the numerator and denominator follow a gamma distribution. Furthermore, the sums in the latter expression are independent, and, after scaling, we find

$$\frac{r-1}{s-r} T_1 = \frac{r-1}{s-r} \frac{\ln(L_s) - \ln(L_r)}{\ln(L_r) - \ln(L_1)} \sim F(2(s-r), 2(r-1)),$$

where $F(\cdot, \cdot)$ denotes an F-distribution with respective degrees of freedom. Now, we can compute a $(1 - \alpha)$ -prediction interval for the s th lower record value by using F-quantiles, namely

$$I_1 = \left[L_r \exp\left(\frac{s-r}{r-1} (\ln(L_r) - \ln(L_1)) F_{1-\frac{\alpha}{2}}(2(s-r), 2(r-1))\right), \right. \\ \left. L_r \exp\left(\frac{s-r}{r-1} (\ln(L_r) - \ln(L_1)) F_{\frac{\alpha}{2}}(2(s-r), 2(r-1))\right) \right],$$

where $\alpha \in (0, 1)$. Next, we consider

$$T_2 = \frac{\ln(L_s) - \ln(L_r)}{\ln(L_r) - \ln(\lambda)} \stackrel{d}{=} \frac{\sum_{i=r+1}^s Z_i}{\sum_{i=1}^r Z_i}.$$

Thus, we find

$$\frac{r}{s-r} \frac{\ln(L_s) - \ln(L_r)}{\ln(L_r) - \ln(\lambda)} \sim F(2(s-r), 2r).$$

Hence, a $(1 - \alpha)$ -prediction interval is given by

$$I_2 = \left[L_r \exp\left(\frac{s-r}{r} (\ln(L_r) - \ln(\lambda)) F_{1-\frac{\alpha}{2}}(2(s-r), 2r)\right), \right. \\ \left. L_r \exp\left(\frac{s-r}{r} (\ln(L_r) - \ln(\lambda)) F_{\frac{\alpha}{2}}(2(s-r), 2r)\right) \right].$$

A generalized version of the statistic T_2 is used in Wang et al. [20] in a setting with unknown parameter λ . For a known parameter λ , prediction interval I_2 will be applied instead of I_1 .

In addition to those exact intervals, we study some approximate $(1 - \alpha)$ -prediction intervals based on statistics similar to T_1 and T_2 , starting with $T_3 = \ln(L_r) - \ln(L_s)$. It follows that

$$2\beta T_3 \stackrel{d}{=} 2\beta \frac{1}{\beta} \sum_{i=r+1}^s Z_i \sim \chi^2(2(s-r)),$$

where $\chi^2(\cdot)$ denotes a χ^2 -distribution with respective degrees of freedom. Plugging in $\hat{\beta}_{ML} = r/(\ln(\lambda) - \ln(L_r))$ for β leads to the approximate $(1 - \alpha)$ -prediction interval

$$I_3 = \left[L_r \exp\left(\frac{\ln(L_r) - \ln(\lambda)}{2r} \chi_{1-\frac{\alpha}{2}}^2(2(s-r))\right), L_r \exp\left(\frac{\ln(L_r) - \ln(\lambda)}{2r} \chi_{\frac{\alpha}{2}}^2(2(s-r))\right) \right],$$

containing χ^2 -quantiles. Furthermore, we consider the statistic $T_4 = \ln(\lambda) - \ln(L_s)$. From Equation (3), one gets

$$2\beta T_4 \stackrel{d}{=} 2 \sum_{i=1}^s Z_i \sim \chi^2(2s)$$

and plugging in $\hat{\beta}_{ML}$, leads to

$$I_4 = \left[\lambda \exp\left(\frac{\ln(L_r) - \ln(\lambda)}{2r} \chi_{1-\frac{\alpha}{2}}^2(2s)\right), \lambda \exp\left(\frac{\ln(L_r) - \ln(\lambda)}{2r} \chi_{\frac{\alpha}{2}}^2(2s)\right) \right],$$

as an approximate $(1 - \alpha)$ -prediction interval.

The prediction intervals I_1, \dots, I_4 are compared via the expected length and coverage percentage in a simulation study below.

2.2.2. Upper Record Values from Pareto Distributions

The corresponding prediction intervals for upper record values based on Pareto distributed random variables can be obtained analogously and are given by

$$\begin{aligned} I'_1 &= \left[R_r \exp\left(\frac{s-r}{r-1}(\ln(R_r) - \ln(R_1))F_{\frac{\alpha}{2}}(2(s-r), 2(r-1))\right), \right. \\ &\quad \left. R_r \exp\left(\frac{s-r}{r-1}(\ln(R_r) - \ln(R_1))F_{1-\frac{\alpha}{2}}(2(s-r), 2(r-1))\right) \right] \\ I'_2 &= \left[R_r \exp\left(\frac{s-r}{r}(\ln(R_r) - \ln(\mu))F_{\frac{\alpha}{2}}(2(s-r), 2r)\right), \right. \\ &\quad \left. R_r \exp\left(\frac{s-r}{r}(\ln(R_r) - \ln(\mu))F_{1-\frac{\alpha}{2}}(2(s-r), 2r)\right) \right] \\ I'_3 &= \left[R_r \exp\left(\frac{\ln(R_r) - \ln(\mu)}{2r}\chi_{\frac{\alpha}{2}}^2(2(s-r))\right), R_r \exp\left(\frac{\ln(R_r) - \ln(\mu)}{2r}\chi_{1-\frac{\alpha}{2}}^2(2(s-r))\right) \right] \\ I'_4 &= \left[\mu \exp\left(\frac{\ln(R_r) - \ln(\mu)}{2r}\chi_{\frac{\alpha}{2}}^2(2s)\right), \mu \exp\left(\frac{\ln(R_r) - \ln(\mu)}{2r}\chi_{1-\frac{\alpha}{2}}^2(2s)\right) \right]. \end{aligned}$$

I'_1 was introduced in Asgharzadeh et al. [22]. I'_2 and I'_4 can be derived from the corresponding intervals in Awad and Raqab [21] via an analogous formulation of the transformation in Equation (6) for upper record values. Raqab et al. [15] proposed I'_3 .

2.2.3. Expected Lengths of Prediction Intervals

For the expected lengths $l_i = E(\hat{b}_i(L_*) - \hat{a}_i(L_*))$ of the prediction intervals $I_i = [\hat{a}_i(L_*), \hat{b}_i(L_*)]$, $i = 1, \dots, 4$, for future lower record values, the following expressions can be derived.

According to Equation (3), the expectation of the upper bound of I_1 is given by

$$\begin{aligned} &E\left(L_r \exp\left(\frac{s-r}{r-1}(\ln(L_r) - \ln(L_1))F_{\frac{\alpha}{2}}(2(s-r), 2(r-1))\right)\right) \\ &= E\left(\exp\left(\ln(L_r) + \frac{s-r}{r-1}F_{\frac{\alpha}{2}}(2(s-r), 2(r-1))(\ln(L_r) - \ln(L_1))\right)\right) \\ &= E\left(\exp\left(\ln(\lambda) - \frac{1}{\beta} \sum_{i=1}^r Z_i + \frac{s-r}{r-1}F_{\frac{\alpha}{2}}(2(s-r), 2(r-1))\left(-\frac{1}{\beta} \sum_{i=1}^r Z_i + \frac{1}{\beta} Z_1\right)\right)\right) \\ &= \lambda E\left(\exp\left(-\frac{1}{\beta} Z_1\right)\right) \prod_{i=2}^r E\left(\exp\left(\frac{1}{\beta} \left(-1 - \frac{s-r}{r-1}F_{\frac{\alpha}{2}}(2(s-r), 2(r-1))\right) Z_i\right)\right) \\ &= \lambda \frac{1}{1 + \frac{1}{\beta}} \left(\frac{1}{1 + \left(1 + \frac{s-r}{r-1}F_{\frac{\alpha}{2}}(2(s-r), 2(r-1))\right) \frac{1}{\beta}} \right)^{r-1}. \end{aligned}$$

In the last step, it was used that the moment-generating function of a standard exponential random variable Z takes the form $E(\exp(tZ)) = \frac{1}{1-t}$ for $t < 1$. Hence,

$$\begin{aligned} l_1 &= \frac{\lambda((r-1)\beta)^{r-1}}{1 + \frac{1}{\beta}} \left(\left(\frac{1}{(r-1)\beta + r - 1 + (s-r)F_{\frac{\alpha}{2}}(2(s-r), 2(r-1))} \right)^{r-1} \right. \\ &\quad \left. - \left(\frac{1}{(r-1)\beta + r - 1 + (s-r)F_{1-\frac{\alpha}{2}}(2(s-r), 2(r-1))} \right)^{r-1} \right). \end{aligned}$$

The expected lengths of the other prediction intervals can be calculated similarly. They are given by

$$l_2 = \lambda(r\beta)^r \left(\left(\frac{1}{\beta r + r + (s-r)F_{\frac{\alpha}{2}}(2(s-r), 2r)} \right)^r - \left(\frac{1}{\beta r + r + (s-r)F_{1-\frac{\alpha}{2}}(2(s-r), 2r)} \right)^r \right)$$

$$l_3 = \lambda(2r\beta)^r \left(\left(\frac{1}{2\beta r + 2r + \chi_{\frac{\alpha}{2}}^2(2(s-r))} \right)^r - \left(\frac{1}{2\beta r + 2r + \chi_{1-\frac{\alpha}{2}}^2(2(s-r))} \right)^r \right)$$

$$l_4 = \lambda(2r\beta)^r \left(\left(\frac{1}{2\beta r + \chi_{\frac{\alpha}{2}}^2(2s)} \right)^r - \left(\frac{1}{2\beta r + \chi_{1-\frac{\alpha}{2}}^2(2s)} \right)^r \right).$$

In the comparisons of prediction intervals, we restricted ourselves to the most relevant cases in practice, namely $s = r + 1$ and $s = r + 2$. Moreover, as examples, we chose the parameters as $(\lambda, \beta) = (11.3, 70)$ and $(\lambda, \beta) = (45.5, 90)$, which corresponded to the women's 100 m and the men's 400 m, respectively (see Tables 1 and 2 and Section 3).

Table 1. Expected lengths of the 95% and 90% prediction intervals I_1, \dots, I_4 for the s th lower record value based on the first r records from a $Pow(\lambda, \beta)$ distribution.

	$\alpha = 5\%$				$\alpha = 10\%$			
	I_1	I_2	I_3	I_4	I_1	I_2	I_3	I_4
$\lambda = 11.3, \beta = 70$								
$r = 3, s = 4$	1.45	1.03	0.54	1.13	0.98	0.74	0.44	0.95
$r = 8, s = 9$	0.66	0.64	0.51	1.61	0.51	0.49	0.41	1.35
$r = 8, s = 10$	1.01	0.97	0.72	1.67	0.79	0.77	0.60	1.40
$r = 25, s = 26$	0.43	0.43	0.40	2.17	0.34	0.34	0.32	1.83
$r = 25, s = 27$	0.63	0.63	0.57	2.18	0.51	0.51	0.47	1.83
$r = 25, s = 28$	0.78	0.78	0.70	2.19	0.64	0.64	0.58	1.84
$\lambda = 45.5, \beta = 90$								
$r = 3, s = 4$	4.72	3.32	1.72	3.61	3.15	2.37	1.39	3.02
$r = 8, s = 9$	2.15	2.07	1.64	5.21	1.65	1.60	1.32	4.38
$r = 8, s = 10$	3.29	3.16	2.35	5.43	2.59	2.50	1.95	4.56
$r = 25, s = 26$	1.47	1.46	1.36	7.41	1.17	1.17	1.10	6.23
$r = 25, s = 27$	2.16	2.15	1.96	7.46	1.76	1.75	1.62	6.28
$r = 25, s = 28$	2.69	2.68	2.40	7.52	2.21	2.20	2.00	6.32

Table 2. Percentages of coverage of the 90% prediction intervals I_1, \dots, I_4 for the s th lower record value based on $n = 10,000$ sequences of the first r lower record values from a $Pow(\lambda, \beta)$ distribution.

	I_1	I_2	I_3	I_4
$\lambda = 11.3, \beta = 70$				
$r = 3, s = 4$	0.9041	0.9001	0.8229	0.9404
$r = 8, s = 9$	0.8965	0.8989	0.8716	0.9910
$r = 8, s = 10$	0.8995	0.8994	0.8492	0.9785
$r = 25, s = 26$	0.8995	0.8985	0.8890	0.9998
$r = 25, s = 27$	0.8980	0.8990	0.8832	0.9988
$r = 25, s = 28$	0.9032	0.9040	0.8808	0.9971

Table 2. Cont.

	I_1	I_2	I_3	I_4
$\lambda = 45.5, \beta = 90$				
$r = 3, s = 4$	0.9041	0.9001	0.8229	0.9404
$r = 8, s = 9$	0.8965	0.8989	0.8716	0.9910
$r = 8, s = 10$	0.8995	0.8994	0.8492	0.9785
$r = 25, s = 26$	0.8995	0.8985	0.8890	0.9998
$r = 25, s = 27$	0.8980	0.8990	0.8832	0.9988
$r = 25, s = 28$	0.9032	0.9040	0.8808	0.9971

In Table 1, the expected lengths of the prediction intervals I_1, \dots, I_4 are shown for the two chosen parameter combinations and for different values of r and s . The prediction interval I_2 which uses λ as a known parameter outperforms I_1 , where the first record value is used. Compared with the other intervals, I_4 has large expected lengths, which are even increasing in the number of record values used in the prediction method. The asymptotic method I_3 tends to have slightly shorter prediction intervals than the exact ones, I_1 and I_2 . Figures 1 and 2 illustrate the expected lengths of all intervals for $s = r + 1$ and $s = r + 2$, respectively.

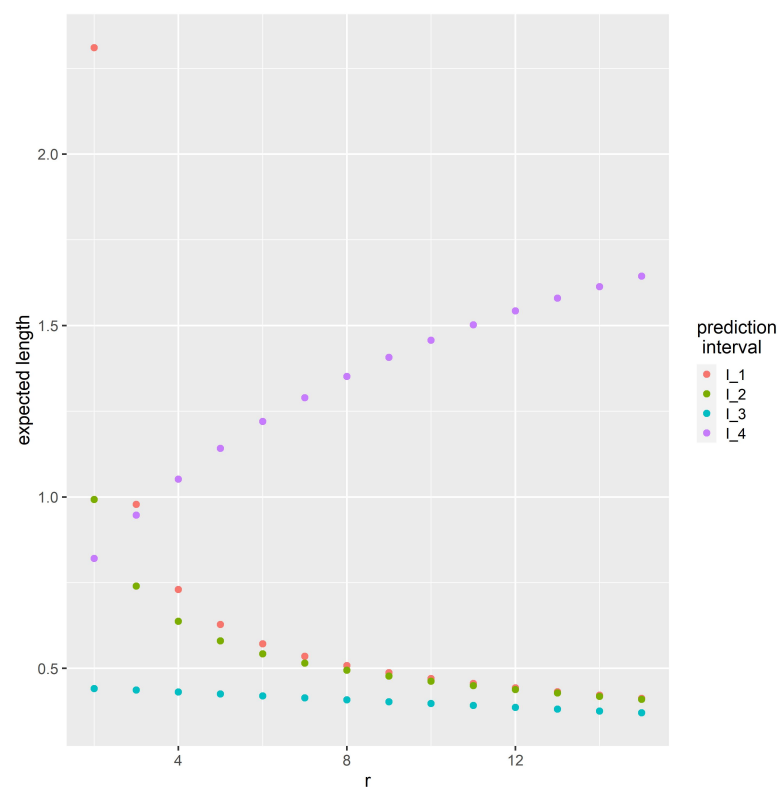


Figure 1. Expected lengths of the 90% prediction intervals I_1, \dots, I_4 for the next lower record value ($s = r + 1$) based on the first r lower record values from a $Pow(\lambda, \beta)$ distribution with $\lambda = 11.3$ and $\beta = 70$.

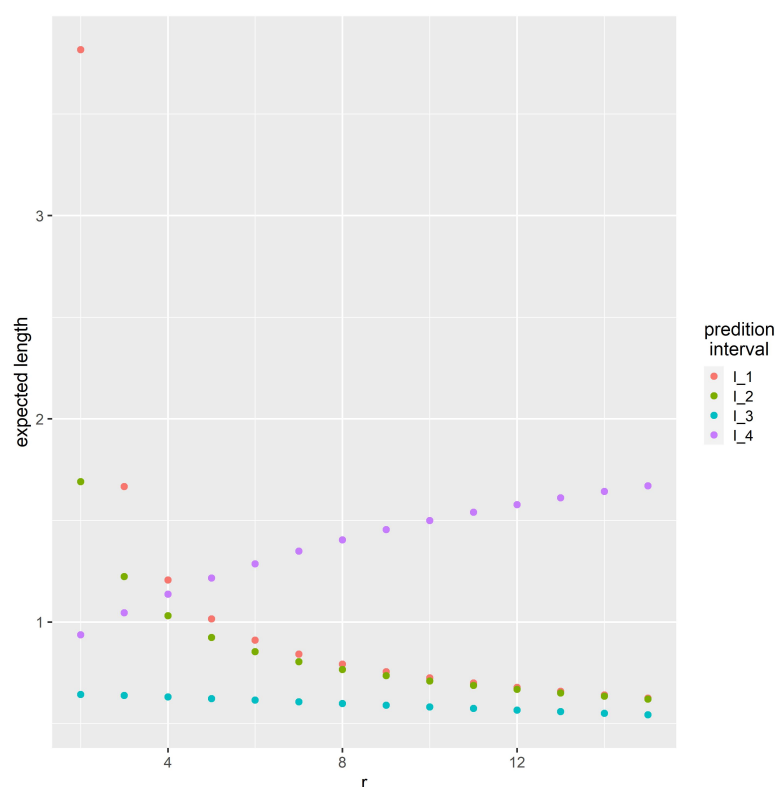


Figure 2. Expected lengths of the 90% prediction intervals I_1, \dots, I_4 for the next but one lower record value ($s = r + 2$) based on the first r lower record values from a $Pow(\lambda, \beta)$ distribution with $\lambda = 11.3$ and $\beta = 70$.

We performed a simulation study concerning the percentages of coverage of the prediction intervals. We generated $n = 10,000$ sequences of record values from a power function distribution and considered the first r record values as known and the remaining records as unknown. Then, we computed the four prediction intervals based on these sequences and their empirical percentages of coverage. The results in Table 2 show that the empirical coverage of the exact prediction intervals I_1 and I_2 was always close to $1 - \alpha$. As observed in Table 1, I_3 had small expected lengths, which corresponded to percentages of coverage (considerably) smaller than 90%, throughout. The asymptotic prediction interval I_4 turned out to be too conservative. Summarizing, the exact prediction intervals I_1 and I_2 seemed to outperform the approximate approaches I_3 and I_4 .

3. Application to Athletics and American Football Data

Records play an important role in athletics. We applied statistical prediction methods for future lower and upper record values, namely, future world records, in athletic events, and we focused on the respective next record, i.e., $s = r + 1$ was chosen. We assumed that lower records, as in running events, were based on a power function distribution, and that upper records, as in throwing and jumping events, were based on a Pareto distribution. These assumptions have to be justified in a given data situation. As an example, we considered the women's 100 m; data of this discipline have been studied by extreme value methods before (Einmahl and Magnus [2], Einmahl and Smeets [4], Stephenson and Tawn [7]). We used the data provided at <https://www.worldathletics.org/records/all-time-toplists>, accessed on 19 December 2019. These data contain the personal bests of respective athletes, and we considered results until 2018. Since the times were taken with an accuracy of 0.01 s, we further adjusted those times in order to eliminate ties as in Einmahl and Magnus [2]. A histogram of these slightly modified times is given in Figure 3.

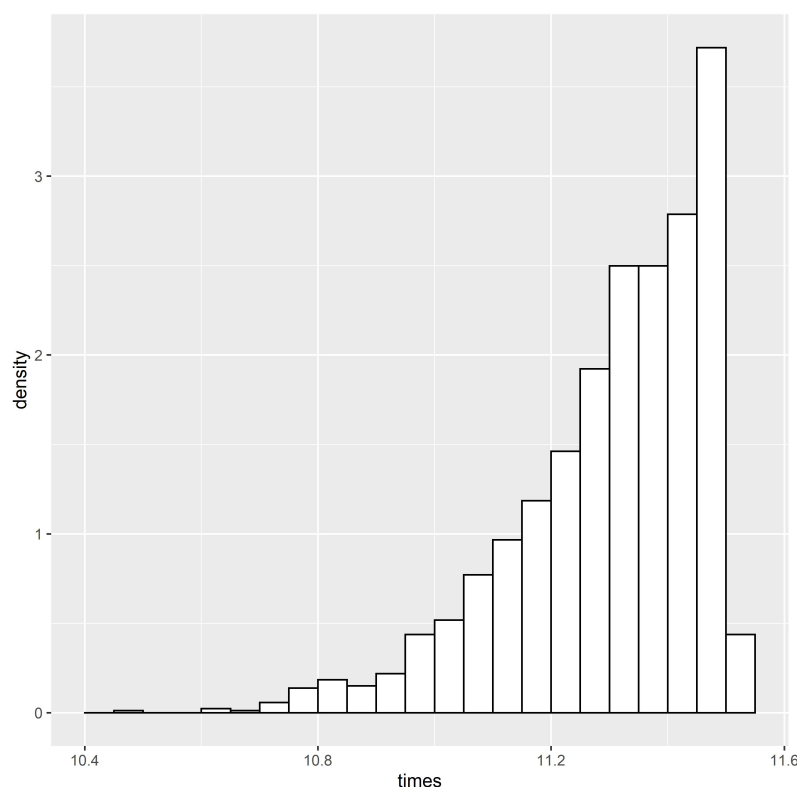


Figure 3. Histogram of the times of women's 100 m in seconds.

The common record model as introduced in Section 2 is based on a sequence of iid random variables. While the assumption of independence seems to be reasonable, the assumption of identical distribution is questionable due to a possible trend in the data over time. In what follows, we stick to the iid situation. For the reason of comparability of sports results and in order to better approximately meet the assumption of identical distributions, we considered the times of top athletes only, whose running times were below a given threshold. In Figure 4 and in our prediction results, this threshold was chosen to be $\lambda = 11.3$.

The curve in Figure 4 illustrates the density function of a power function distribution with parameters $\lambda = 11.3$ and $\beta = 69.73$, where the latter value is a maximum likelihood estimate. Figure 4 suggests that the power function assumption seems to be reasonable. However, the respective quantile–quantile (Q–Q) plot in Figure 5 gives rise to question the assumptions.

Maybe the power function assumption should be modified as the smallest values indicate, or the assumptions of independence and identical distributions should be modified. Nevertheless, due to the histogram plot and for illustration of our theoretical results, we stuck to the power function assumption.

Then, we applied the point predictor MPSP for lower and upper records and the prediction intervals I_1/I'_1 , I_2/I'_2 and I_3/I'_3 , which were shown in Section 2. The world record progression for the different events was recorded in <https://www.worldathletics.org/records/by-category/world-records>, accessed on 25 November 2022. Unfortunately, not all of these lists were consistent. For example, some times in the running events were measured by hand and therefore had an accuracy of 0.1 instead of 0.01 for electronically measured times. In this listing, a record of 11.20 was regarded as faster than the old record 11.1 for the women's 100 m. To avoid such data problems, we excluded times measured by hand. To illustrate the prediction results, Table 3 shows world records in the women's 100 m along with respective MPSP and prediction intervals I_1 , I_2 and I_3 based on the previous record values. The point predictor (MPSP) and the prediction interval shown in row i of Table 3 are based on the previous records listed in rows $1, \dots, i - 1$, $i \geq 2$.

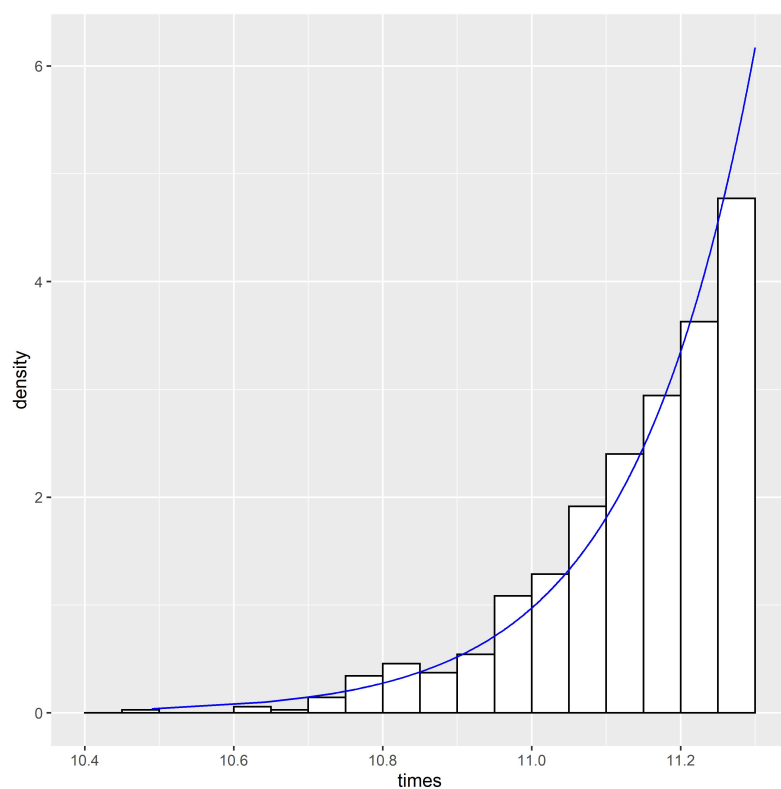


Figure 4. Histogram of the results of women's 100 m faster than 11.3 s and pdf of the power function distribution with $\lambda = 11.3$ and $\beta = 69.73$, where the latter is the maximum likelihood estimate based on all times below 11.3 as upper threshold.

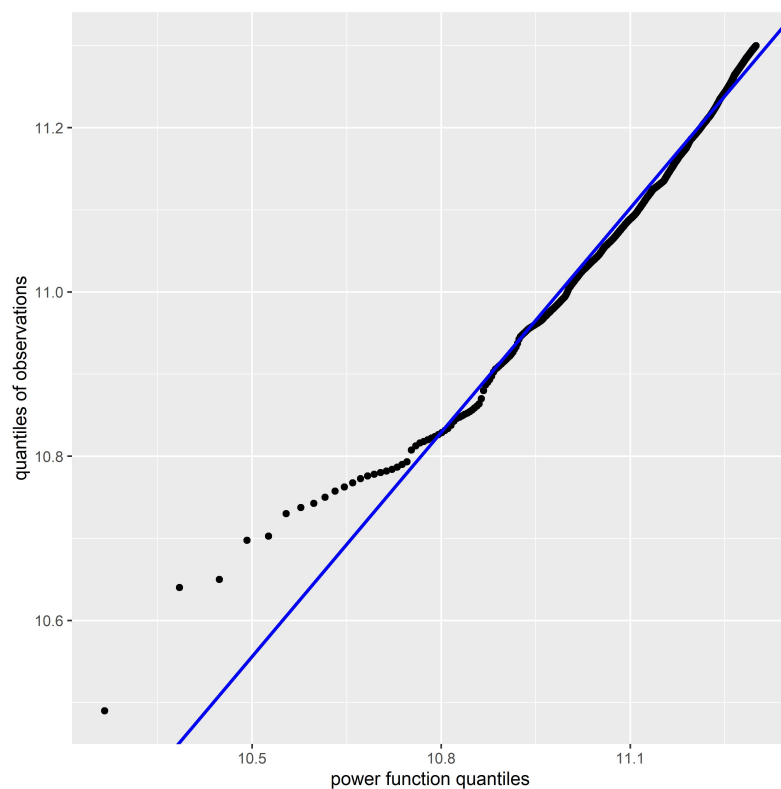


Figure 5. Power function Q–Q plot for the times of women's 100 m faster than 11.3 s, where the five smallest times were omitted when calculating the regression line.

Table 3. World records, maximum product of spacings predictor and (approximate) 90% prediction intervals I_1 , I_2 and I_3 for the next record of the women's 100 m based on the previous records.

s	World Record	MPSP	I_1	I_2	I_3
1	11.20				
2	11.08	11.10		[9.46, 11.19]	[10.91, 11.19]
3	11.07	10.97	[9.03, 11.07]	[10.35, 11.07]	[10.76, 11.07]
4	11.04	10.99	[10.63, 11.07]	[10.69, 11.07]	[10.84, 11.07]
5	11.01	10.98	[10.77, 11.04]	[10.76, 11.04]	[10.85, 11.04]
6	10.88	10.95	[10.80, 11.01]	[10.78, 11.01]	[10.84, 11.01]
7	10.81	10.81	[10.62, 10.88]	[10.62, 10.88]	[10.68, 10.88]
8	10.79	10.74	[10.56, 10.81]	[10.56, 10.81]	[10.61, 10.81]
9	10.76	10.73	[10.58, 10.79]	[10.57, 10.79]	[10.61, 10.79]
10	10.49	10.70	[10.57, 10.76]	[10.55, 10.76]	[10.59, 10.76]
11		10.41	[10.22, 10.49]	[10.22, 10.49]	[10.26, 10.49]

It can be observed that most world records were close to the MPSP and were within the prediction intervals. An exception was the record ran by Florence Griffith-Joyner in 1988. It was considerably smaller than the MPSP and the lower bound of the statistical prediction intervals. Moreover, such an exceptional record had a strong effect on the statistical prediction of subsequent record values.

By using the same procedure in other running events and an analogous approach for upper records in throwing and jumping events with underlying Pareto distributions, where the distributional assumption was approximately met, we derived the results shown in Table 4.

Table 4. World records (until 2022), maximum product of spacings predictor and 90% prediction interval I_2/I_2' for the next record based on the previous records for various athletic events.

Event	Women				Men			
	Record	r	MPSP	Prediction Interval	Record	r	MPSP	Prediction Interval
100 m	10.49	10	10.41	[10.22, 10.49]	9.58	13	9.53	[9.40, 9.58]
100/110 m hurdles	12.12	9	12.04	[11.83, 12.12]	12.80	9	12.71	[12.50, 12.80]
200 m	21.34	9	21.15	[20.68, 21.33]	19.19	5	18.90	[18.03, 19.18]
400 m	47.60	12	47.25	[46.42, 47.58]	43.03	4	42.46	[40.53, 43.00]
800 m	1:53.28	2	1:50.06	[1:32.74, 1:53.11]	1:40.91	7	1:40.20	[1:38.29, 1:40.87]
1500 m	3:50.07	3	3:45.61	[3:28.01, 3:49.84]	3:26.00	8	3:24.55	[3:20.77, 3:25.92]
10,000 m	29:01.03	9	28:40.22	[27:48.19, 28:59.95]	26:11.00	13	26:02.20	[25:41.55, 26:10.55]
Marathon	2:14:04	2	2:06:45	[1:30:47, 2:13:41]	2:01:09	8	2:00:11	[1:57:40, 2:01:06]
Shot put	22.63	26	22.81	[22.64, 23.19]	23.37	16	23.56	[23.38, 24.01]
Javelin throw	72.28	3	78.70	[72.60, 111.95]	98.48	8	101.07	[98.61, 108.23]
Discus throw	76.80	17	77.56	[76.84, 79.31]	74.08	12	74.89	[74.12, 76.88]
Long jump	7.52	14	7.57	[7.52, 7.70]	8.95	9	9.06	[8.96, 9.36]
High jump	2.09	13	2.10	[2.09, 2.13]	2.45	22	2.46	[2.45, 2.48]

Obviously, the number of world records was small in the women's 800 m, 1500 m, marathon and javelin throw. Therefore, the predictions in Table 4 may differ from the actual next record. Moreover, the respective prediction intervals were quite large. For example, the lower bound for the next marathon record was nonsatisfying.

However, in the men's marathon, the lower bound was smaller than two hours as well. The number of records was $r = 8$, so the prediction interval was expected to be quite reasonable having a lower bound below the two-hour mark. In October 2019, the world record holder Eliud Kipchoge broke this mark under special conditions regarding, for example, the course, wherefore it was not considered an actual world record.

In general, prediction intervals tend to be large, if the number of observed records is fairly small. Thus, in such a real data situation, the prediction interval may not be of practical use (see Table 4).

In addition to the analysis of athletic events, we applied point and interval prediction to data from American football. A possible data set could result from the NFL Combine with its athletics disciplines.

However, we focused on data resulting from actual football games in the NFL. To combine most relevant metrics in the analysis of a player's performance on the field, we considered so-called fantasy football points (see, e.g., <https://fantasy.nfl.com/research/scoringleaders>, accessed on 2 March 2021). These fantasy points are calculated according to the points per reception (PPR) scoring scheme (see Table 5 and the glossary at <https://www.pro-football-reference.com>, accessed on 14 January 2020), for each game separately. Many providers of football data additionally offer fantasy points of players that are added up to a full season performance. Here, we focus on single game performances.

Table 5. List of points given in fantasy football's PPR scoring.

Event	Score
Passing	1 point per 25 yards
Passing touchdowns	4 points
Interceptions thrown	−2 points
Rushing/receiving yards	1 point per 10 yards
Receptions	1 point
Touchdowns	6 points
2-Point conversions	2 points
Fumbles lost	−2 points

Related to these data of fantasy football points, we may define world records in American football. When measured by means of the PPR scheme, the respective world record performance (as of December 2019) by quarterback Michael Vick was composed of 333 passing yards, 4 passing touchdowns, 80 rushing yards and 2 rushing touchdowns in a single game, which yielded

$$\frac{333}{25} + 4 \times 4 + \frac{80}{10} + 2 \times 6 = 49.32 \text{ fantasy points.}$$

We distinguished between the skill positions quarterback, wide receiver, running back and tight end. Other players such as linemen, defense players and special teams were not included in the analysis. The points of kickers and whole defense teams were measured in the standard scoring scheme as well, but they followed a discrete distribution and were therefore not considered here.

We determined the sequences of records for quarterbacks, wide receivers, running backs and tight ends as the basis for point and interval prediction. The underlying game data are provided at <https://www.pro-football-reference.com> (accessed on 14 January 2020, publicly available until 2020) and cover each game of the regular seasons from 2000 to 2019. The fantasy football points were rounded to one digit for quarterbacks as well. The Pareto assumption could be justified by comparing the histogram and the estimated density function as above. Only for the tight end's points did a Pareto distribution not seem to be reasonable. Since there were just two records in the wide receiver's data, we did not compute predictions for this position either. The progressions of the world records and the corresponding predictions can be found in Tables 6 and 7.

Table 6. World records (2000–2019), maximum product of spacings predictor and (approximate) 90% prediction intervals I'_1 , I'_2 and I'_3 for the next record based on the previous records for the fantasy points of quarterbacks.

s	Player	Team	Records	MPSP	I'_1	I'_2	I'_3
1	Cade McNown	CHI	34.3				
2	Trent Green	STL	36.3	42.0		[34.7, 1621.4]	[34.7, 63.0]
3	Peyton Manning	IND	37.4	41.3	[36.4, 106.5]	[36.5, 89.4]	[36.5, 53.6]
4	Trent Green	KAN	37.9	41.2	[37.5, 50.5]	[37.6, 61.4]	[37.6, 49.9]
5	Michael Vick	ATL	38.2	40.9	[38.0, 45.0]	[38.0, 53.1]	[38.0, 47.5]
6	Daunte Culpepper	MIN	41.8	40.6	[38.3, 43.1]	[38.3, 49.3]	[38.3, 46.0]
7	Michael Vick	PHI	49.3	44.7	[41.9, 49.2]	[41.9, 54.2]	[41.9, 51.1]
8				53.4	[49.5, 62.4]	[49.5, 66.7]	[49.5, 62.8]

Table 7. World records (2000–2019), maximum product of spacings predictor and (approximate) 90% prediction intervals I'_1 , I'_2 and I'_3 for the next record based on the previous records for the fantasy points of running backs.

s	Player	Team	Records	MPSP	I'_1	I'_2	I'_3
1	Duce Staley	PHI	36.2				
2	Marshall Faulk	STL	44.9	43.7		[36.6, 1284.9]	[36.6, 63.6]
3	Marshall Faulk	STL	45.6	54.9	[45.4, 2688.2]	[45.4, 182.1]	[45.4, 82.1]
4	Fred Taylor	JAX	51.8	52.4	[45.9, 101.6]	[45.9, 93.5]	[45.9, 69.3]
5	Shaun Alexander	SEA	56.1	59.4	[52.1, 95.7]	[52.2, 95.2]	[52.2, 78.0]
6	Clinton Portis	DEN	57.4	63.6	[56.4, 91.4]	[56.5, 93.8]	[56.5, 81.6]
7	Jamaal Charles	KAN	59.5	64.0	[57.7, 83.8]	[57.7, 87.4]	[57.7, 79.4]
8				65.6	[59.8, 82.1]	[59.8, 85.8]	[59.8, 79.8]

Using box score statistics such as fantasy points to measure a player's performance has been discussed intensively, since it does not weight the yardage made by its effectiveness in a drive. Four yards at third and three are regarded as valuable as four yards at third and five, although the outcome of the play is different.

4. Discussion

We developed and applied point and interval prediction based on a sequence of lower or upper record values to various data sets from athletics and American football to come up with predictions of future record values. In the examples, these predicted values were successively compared with actual subsequent record values. Several prediction methods were discussed and results for lower record values were established, which were used, e.g., for analyzing data from running disciplines in athletics. The procedure as well as the results can also be applied to other sports disciplines. In the model of common record values, we assumed that the considered record values were smallest or largest observations of a sequence of independent and identically distributed random variables. In particular, this assumption of underlying identical distributions presumed that the athletes were able to perform on the same level. In order to approximately meet the assumption of identical distribution, we only considered results better than some threshold interpreted in such a way that we considered just top athletes and top performances. Here, we applied power function distributions with respect to lower record values and Pareto distributions with respect to upper record values, which often are seen to have a reasonable fit to the data. Of course, there are disciplines mentioned in this article, where the distributional assumptions are not really justified, such as in women's 800 m and in tight end's fantasy points. Moreover, certain dependencies within sports data cannot be avoided. For example, some sportspersons in athletics broke a record more than once and therefore, their names occur several times in the list of records. These results cannot be regarded as independent. However, if there are many people involved in a list of records, independence may be assumed at least approximately. Furthermore, in sports such as American football, there are dependencies among a team or between competitors. Despite some possible criticism on the model of common record values in sports applications, the prediction methods performed quite well. Most world records and point predictions lay within the respective prediction intervals. Prominent exceptions are, e.g., the records of Florence Griffith-Joyner in the women's 100 m and Bob Beamon's world record in the men's long jump. Such exceptional performances were much better than expected and had a strong influence on statistical predictions.

5. Conclusions

Based on successive upper or lower record values in a sequence of independent and identically distributed random variables with an underlying Pareto distribution and a power function distribution, respectively, we obtained explicit statistical methods for point prediction and interval prediction. In this article, the procedures were derived in the case of lower records. Prediction intervals were compared by the criterion of expected lengths, and percentages of coverage were tabulated. The results were illustrated via real data

sets from several athletics disciplines and American football. The forecasting of records has been considered in the literature before, mainly by means of extreme value theory, whereas we focused on statistical prediction methods to come up with point and interval predictions. When applied to the data sets, the model assumptions of independence and identical distribution of random variables as well as the specification of the underlying distribution have to be discussed in view of the developments in the respective sports disciplines and the prediction results. This may lead one to consider other underlying distributions providing a better fit as well as to refine the record model by incorporating a possible trend in the data over time. A heuristic approach to take into account a trend in the underlying data could be to estimate this trend, to rescale the original data ahead of prediction and then to modify the prediction results accordingly. Prediction within a thorough statistical model will be the subject matter of our further research.

Author Contributions: Conceptualization, C.E. and U.K.; methodology, U.K. and G.V.; software, C.E.; validation, C.E., U.K. and G.V.; formal analysis, C.E.; investigation, C.E. and G.V.; data curation, C.E.; writing—original draft preparation, C.E. and G.V.; writing—review and editing, U.K.; visualization, C.E. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Publicly available datasets were analyzed in this study. These data can be found here: <https://www.worldathletics.org/records/all-time-toplists> (accessed on 19 December 2019), <https://www.worldathletics.org/records/by-category/world-records> (accessed on 25 November 2022), <https://www.pro-football-reference.com> (accessed on 14 January 2020). The data from <https://www.pro-football-reference.com> were publicly available until 2020.

Acknowledgments: The authors would like to thank the reviewers for their helpful comments and suggestions.

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

Abbreviations

The following abbreviations are used in this manuscript:

iid	independent and identically distributed
cdf	cumulative distribution function
pdf	probability density function
MLP	maximum likelihood predictor
MOLP	maximum observed likelihood predictor
MPSP	maximum product of spacings predictor
QQ	quantile–quantile
NFL	National Football League
PPR	points per reception

References

1. Noubary, R.D. A procedure for prediction of sports records. *J. Quant. Anal. Sport.* **2005**, *1*, [CrossRef]
2. Einmahl, J.H.J.; Magnus, J.R. Records in athletics through extreme-value theory. *J. Am. Stat. Assoc.* **2008**, *103*, 1382–1391. [CrossRef]
3. Noubary, R.D. Tail modeling, track and field records, and Bolt’s effect. *J. Quant. Anal. Sport.* **2010**, *6*, [CrossRef]
4. Einmahl, J.H.J.; Smeets, S.G.W.R. Ultimate 100-m world records through extreme-value theory. *Stat. Neerl.* **2011**, *65*, 32–42. [CrossRef]
5. Henriques-Rodrigues, L.; Gomes, M.; Pestana, D. Statistics of extremes in athletics. *Revstat Stat. J.* **2011**, *9*, 127–153. [CrossRef]
6. Fraga Alves, I.; de Haan, L.; Neves, C. How far can man go? In *Advances in Theoretical and Applied Statistics*; Torelli, N., Pesarin, F., Bar-Hen, A., Eds.; Springer: Heidelberg, Germany, 2013; pp. 187–197. [CrossRef]
7. Stephenson, A.G.; Tawn, J.A. Determining the best track performances of all time using a conceptual population model for athletics records. *J. Quant. Anal. Sport.* **2013**, *9*, 67–76. [CrossRef]

8. Adam, M.B.; Tawn, J.A. Modelling record times in sport with extreme value methods. *Malays. J. Math. Sci.* **2016**, *10*, 1–21.
9. Albert, J.; Glickman, M.E.; Swartz, T.B.; Koning, R.H. (Eds.) *Handbook of Statistical Methods and Analyses in Sports*; CRC Press: Boca Raton, FL, USA, 2019.
10. Wunderlich, F.; Memmert, D. Forecasting the outcomes of sports events: A review. *Eur. J. Sport Sci.* **2021**, *21*, 944–957. [[CrossRef](#)] [[PubMed](#)]
11. Arnold, B.C.; Balakrishnan, N.; Nagaraja, H.N. *Records*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 1998. [[CrossRef](#)]
12. Kaminsky, K.S.; Rhodin, L.S. Maximum likelihood prediction. *Ann. Inst. Stat. Math.* **1985**, *37*, 507–517. [[CrossRef](#)]
13. Volovskiy, G.; Kamps, U. Maximum observed likelihood prediction of future record values. *TEST* **2020**, *29*, 1072–1097. [[CrossRef](#)]
14. Volovskiy, G.; Kamps, U. Maximum product of spacings prediction of future record values. *Metrika* **2020**, *83*, 853–868. [[CrossRef](#)]
15. Raqab, M.Z.; Ahmadi, J.; Doostparast, M. Statistical inference based on record data from Pareto model. *Statistics* **2007**, *41*, 105–118. [[CrossRef](#)]
16. Ahmadi, J.; Doostparast, M. Bayesian estimation and prediction for some life distributions based on record values. *Stat. Pap.* **2006**, *47*, 373–392. [[CrossRef](#)]
17. Madi, M.T.; Raqab, M.Z. Bayesian prediction of temperature records using the Pareto model. *Environmetrics* **2004**, *15*, 701–710. [[CrossRef](#)]
18. Cheng, R.C.H.; Amin, N.A.K. Estimating parameters in continuous univariate distributions with a shifted origin. *J. R. Stat. Soc. Ser. B* **1983**, *45*, 394–403. [[CrossRef](#)]
19. Ranneby, B. The maximum spacing method. An estimation method related to the maximum likelihood method. *Scand. J. Stat.* **1984**, *11*, 93–112.
20. Wang, B.X.; Yu, K.; Coolen, F.P. Interval estimation for proportional reversed hazard family based on lower record values. *Stat. Probab. Lett.* **2015**, *98*, 115–122. [[CrossRef](#)]
21. Awad, A.M.; Raqab, M.Z. Prediction intervals for the future record values from exponential distribution: Comparative study. *J. Stat. Comput. Simul.* **2000**, *65*, 325–340. [[CrossRef](#)]
22. Asgharzadeh, A.; Abdi, M.; Kuş, C. Interval estimation for the two-parameter Pareto distribution based on record values. *Selçuk J. Appl. Math.* **2011**, 149–161.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.