Statistics of Superior Records

E. Ben-Naim\textsuperscript{1} and P. L. Krapivsky\textsuperscript{2}

\textsuperscript{1}Theoretical Division and Center for Nonlinear Studies, Los Alamos National Laboratory, Los Alamos, New Mexico 87545
\textsuperscript{2}Department of Physics, Boston University, Boston, Massachusetts 02215

We study statistics of records in a sequence of random variables. The running record equals the maximum of all elements in the sequence up to a given point, and we define a superior sequence as one where all running records are above average. We obtain the record distribution for superior sequences in the limit $N \to \infty$ where $N$ is sequence length. Further, we find that the fraction of superior sequences $S_N$ decays algebraically, $S_N \sim N^{-\beta}$. Interestingly, the decay exponent $\beta$ is nontrivial, being the root an integral equation. For example, when the random variables are drawn from a uniform distribution, we find $\beta = 0.450265$. In general, the tail of the distribution function from which the random variables are drawn governs the exponent $\beta$. We also consider the dual question of inferior sequences, where all records are below average, and find that the fraction of inferior sequences $I_N$ decays algebraically, albeit with a different decay exponent, $I_N \sim N^{-\alpha}$.

PACS numbers: 02.50.-r, 05.40.-a, 02.30.Em, 02.50.Cw, 05.45.Tp

Extreme values are an important feature of data sets, and they are widely used to analyze data in fields ranging from engineering \cite{1} to finance \cite{2, 3}. For example, the largest and the smallest data points specify the span of the set. Statistical properties of extreme values play a central role in probability theory and in statistical physics \cite{4–7}. Statistical studies of extreme values typically focus on average and extremal properties of the distribution of extreme values \cite{8, 9}. First passage and persistence properties (see \cite{10, 11} and references therein) enrich our understanding of random processes, yet so far they have not received significant attention in the context of extreme values.

In this study, we investigate first-passage characteristics of extreme values. Specifically, we compare extreme values with their expected average as a measure of “performance”. We track the record, defined as the largest variable in a sequence of uncorrelated random variables, and ask: what is the probability that all records are “superior”, always outperforming the average (Fig. 1). We find that this probability $S_N$ decays algebraically with sequence length $N$ (Fig. 2).

$$S_N \sim N^{-\beta},$$

in the large $N$ limit. Interestingly, the decay exponent $\beta$ is nontrivial, being the root of a transcendental equation. When the random variables are drawn from a uniform distribution with compact support in the unit interval, for which the average record equals $N/(N + 1)$, we find

$$\beta = 0.450265.$$  \hfill (2)

In general, the exponent $\beta$ depends on the tail of the probability distribution function from which the random variables are drawn.

The notion of performance is compelling in climate \cite{12–14}, economics \cite{15–17}, and evolution \cite{18} where records are closely watched. In weather statistics, anomalously hot streaks where the high temperature always exceeds the average high for that day have been studied \cite{10}. In stock markets, it is beneficial to identify companies that are consistently outperforming the average stock index, and in evolution, species with faster-than-average growth rate can be identified.

Consider a set of $N$ independent and identically distributed variables, $\{x_1, x_2, \ldots, x_N\}$. The random variables $0 < x_i < \infty$ are drawn from the probability distribution function $\rho(x)$, with the normalization $\int_0^{\infty} dx \rho(x) = 1$. The cumulative distribution

$$R(x) = \int_x^{\infty} dy \rho(y)$$

(3)

gives the probability of drawing a value larger than $x$, with $R(0) = 1$ and $R(\infty) = 0$. The record $X_N$ equals the largest variable in the series, $X_N = \max(x_1, x_2, \ldots, x_N)$. The probability $R_N(x)$ that the record is larger than $x$ follows immediately from the cumulative distribution

$$R_N(x) = 1 - [1 - R(x)]^N.$$  \hfill (4)

Indeed, the complementary probability that all variables are smaller than $x$, and hence the record is smaller than
there are two possibilities: The \( N \times > A \)
moreover \( F \) \( \text{when all records are above average, that is,} \)
\( X \) with record larger than \( \) as the fraction of superior record sequences of length \( N \).
The distribution \( R \) \( \leq 1 \) for all \( 1 \leq n \leq N \). The quantity is reminiscent of a survival probability \([10]\) since \( R \) is kept finite. Importantly, the scaling function \( \Phi(\alpha) \) determines the parameter \( \alpha \).

By substituting the scaling form \( \Psi(s) \) into the evolution equation \( \Phi(\alpha) = 1 \) where
\[
\alpha = \lim_{N \to \infty} N R N.
\]
We seek a scaling solution
\[
F_N(x) \simeq S_N \Phi(s),
\]
with the scaling variable \( s = R N \) as in \([19]\). By definition, \( F_N(A_N) = S_N \), and hence, \( \Phi(\alpha) = 1 \) where
\[
\alpha = \lim_{N \to \infty} N R N.
\]

This form applies when \( N \to \infty \) and \( R \to 0 \) such that the product \( R N \) is kept finite. Importantly, the scaling function is the same, \( \Phi(s) = 1 - e^{-s} \), for all probability distributions \( \rho(x) \).

We term a record sequence \( (X_1, X_2, \ldots, X_N) \) superior when all records are above average, that is, \( X_n \geq A_n \) for all \( 1 \leq n \leq N \). We are interested in the probability \( S_N \) that a record sequence of length \( N \) is superior. This quantity is reminiscent of a survival probability \([10]\) since we require that the envelope defined by the average is never crossed (Fig. 1).

To find \( S_N \), we have to incorporate the value of the record into our theoretical description. We define \( F_N(x) \) as the fraction of superior record sequences of length \( N \) with record larger than \( x \), namely, \( X_n > x \). The cumulative distribution \( F_N(x) \) is applicable when \( x > A_N \), and moreover \( F_N(A_N) = S_N \) and \( F_N(\infty) = 0 \).

The cumulative distribution obeys the recursion
\[
F_{N+1}(x) = [1 - R(x)] F_N(x) + R(x) S_N
\]
for all \( x > A_{N+1} \). This recursion equation reflects that there are two possibilities: The \((N + 1)\)st element in the sequence may set a new record, or alternatively, the old record may hold. The second term corresponds to the former scenario, and the first term to the latter. The cumulative distribution obeys the difference equation \( F_{N+1} - F_N = R(S_N - F_N) \), where \( F_N \equiv F_N(x) \) and \( R \equiv R(x) \). To determine the asymptotic behavior, we treat \( N \) as a continuous variable, and convert the difference equation \( \Phi(\alpha) = 1 \) into the partial differential equation
\[
\frac{\partial F_N}{\partial N} = R(S_N - F_N).
\]

Essentially, this is as an evolution equation with the sequence length \( N \) playing the role of time.

We seek a scaling solution
\[
F_N(x) \simeq S_N \Phi(s),
\]
with the scaling variable \( s = R N \) as in \([19]\). By definition, \( F_N(A_N) = S_N \), and hence, \( \Phi(\alpha) = 1 \) where
\[
\alpha = \lim_{N \to \infty} N R N.
\]

Here we used the shorthand notation \( R N \equiv R(A_N) \). The cumulative distribution vanishes when \( x \to \infty \), and hence \( \Phi(0) = 0 \). The variable \( s \) varies in the range \( 0 \leq s \leq \alpha \) with the upper bound corresponding to near-average records and the lower bound, to extremely large records. Since \( R \to 0 \) when \( N \to \infty \), equation \( (\Phi) \) shows that the tail of the probability distribution function \( \rho(x) \) determines the parameter \( \alpha \).

By substituting the scaling form \( \Phi(\alpha) = 1 \) into the evolution equation \( \Phi(\alpha) \) and by using the algebraic decay \( \Phi(\alpha) \), we find that the scaling function \( \Phi(s) \) obeys the differential equation
\[
\Phi'(s) + (1 - \beta s^{-1}) \Phi(s) = 1.
\]
We integrate this equation by multiplying both sides by the integrating factor \( s^{-\beta} e^{s} \). Given the boundary condition \( \Phi(0) = 0 \), we obtain \( \Phi(s) = s^{\beta} e^{-s} \int_{0}^{s} du u^{-\beta} e^{u} \), and this expression can be further simplified to
\[
\Phi(s) = s \int_{0}^{1} dz z^{-\beta} e^{s(z-1)}.
\]
By invoking the boundary condition \( \Phi(\alpha) = 1 \), we find the exponent \( \beta \) as root of the transcendental equation
\[
\alpha \int_{0}^{1} dz z^{-\beta} e^{\alpha(z-1)} = 1.
\]
This equation specifies the exponent \( \beta \) and hence, the scaling function \( \Phi(\alpha) \). For arbitrary \( \rho(x) \), the expressions \([13]\) and \([12]\) give the fraction of superior sequences and the record distribution for such sequences. These equations require as input the parameter \( \alpha \) defined in \([10]\) which in turn, requires the average \( A_N \) given in \([19]\). We now apply the general theory above to: (i) the uniform
distribution, (ii) the exponential distribution, and (iii) algebraic distributions, both compact and noncompact.

First, we consider the simplest possible case of a uniform distribution with compact support in a finite interval. Without loss of generality, we chose the unit interval, \( \rho(x) = \begin{cases} 1 & 0 \leq x \leq 1, \\ 0 & x > 1. \end{cases} \) (14)

In this case, the cumulative distribution is linear, \( R(x) = 1 - x \), and from the general equation for the average (5) we immediately obtain

\[ A_N = \frac{N}{N + 1}, \text{ and } \alpha = 1. \] (15)

To obtain \( \alpha \) we simply substitute \( R_N = 1/(N + 1) \) into the definition (10). The integral equation (13) becomes

\[ \int_0^1 dz z^{-\beta} e^{-z} = 1, \] and the root of this equation, which can be obtained with arbitrary precision, is quoted in (2) (see also fig. 2).

The scaling function \( \Phi(s) \) that underlies the cumulative distribution of records (for superior sequences) is shown in figure 3. Also shown is the derivative \( \phi(s) = \Phi'(s) \) that characterizes the record distribution \( f_N = -dF_N/dx \). When \( \rho = 1 \), equation (9) implies the scaling behavior

\[ f_N(x) \simeq N S_N \phi(s) \] (16)

with \( s = N R \). From (11), we obtain \( \phi(0) = 1/(1 - \beta) \) and \( \phi(\alpha) = \beta/\alpha \). The distribution \( \phi(s) \) decreases monotonically with \( s \). One also finds that the average record for a superior sequence, \( \langle x \rangle = \int_0^1 dx f_N(x) x \), behaves as

\[ 1 - \langle x \rangle \simeq C N^{-1}, \quad C = \int_0^1 ds \left[ 1 - \Phi(s) \right]. \] (17)

Since the scaled distribution function \( \phi(s) \) is monotonically decreasing (see Fig. 3) we expect \( C < 1/2 \), and indeed \( C = 0.388476 \). Consequently, the average record is closer to unity than it is to the average \( 1 - A_N \simeq N^{-1} \).

Next, we consider the exponential distribution

\[ \rho(x) = e^{-x}. \] (18)

For a random process where events are independent and occur at constant rate the distribution of waiting time is exponential (19). In this special case, the probability distribution and the cumulative distribution are identical, \( R(x) = \rho(x) \). According to Eq. (5), the average is given by \( A_N = -N \int_0^1 dR (1 - R)^{N-1} \ln R \). By computing this integral, we find the average and the parameter \( \alpha_* \):

\[ A_N = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{N} \quad \text{and} \quad \alpha_* = e^{-\gamma} \] (19)

The parameter \( \alpha_* \) was obtained by substituting the asymptotic behavior \( A_N \simeq \ln N + \gamma \), where \( \gamma = 0.577215 \) is the Euler constant (20), into \( R_N = \exp(-A_N) \) and subsequently, using (10). Plugging \( \alpha_* = 0.561459 \) into the integral equation (13) gives

\[ \beta_* = 0.621127. \] (20)

We also consider a general class of compact distribution functions. Without loss of generality, we assume that the distribution has support in the unit interval \([0 : 1]\). The behavior near the maximum plays a crucial role, and we consider a class of distributions that exhibit algebraic behavior,

\[ \rho(x) \simeq \mu B(1 - x)^{\mu-1}, \] (21)

with \( \mu > 0 \) in the limit \( x \to 1 \). The restriction on \( \mu \) ensures that the distribution \( \rho(x) \) is integrable. From (21),
the cumulative distribution has the asymptotic behavior
\[ R(x) \sim B(1 - x)^\mu \] when \( x \to 1. \) Using the general formulas \([5] \) and \([10] \), we obtain the large-\( N \) asymptotic behavior of the average as well as the parameter \( \alpha \),
\[
1 - A_N \simeq (BN/\alpha)^{-\frac{1}{\alpha}} \quad \text{and} \quad \alpha = \left[ \Gamma(1 + \frac{1}{\mu}) \right]^\mu. \tag{22}
\]
By substituting \( \alpha \) into the integral equation \([13] \), we obtain the exponent \( \beta \). As shown in figure \([4] \) the exponent \( \beta \) varies continuously with \( \mu \) \([21] \). The exponent \( \mu \) parameterizes the shape of the distribution near the maximum. As suggested by equation \([10] \), the tail of the distribution \( \rho(x) \) governs the exponent \( \beta \).

Using the asymptotic behavior \( \Gamma(1 + \epsilon) \simeq 1 - \gamma \epsilon \) for \( \epsilon \to 0 \) with \( \gamma \) the Euler constant, we obtain
\[
\alpha = \left[ \Gamma(1 + \frac{1}{\mu}) \right]^\mu \rightarrow \left[ 1 - \frac{2}{\mu} \right]^\mu \rightarrow e^{-\gamma}
\]
when \( \mu \to \infty \). Hence, the behavior in the limit \( \mu \to \infty \) coincides with that of the exponential distribution. Figure \([4] \) shows that the parameter \( \alpha \) decreases monotonically with \( \mu \) while the exponent \( \beta \) increases monotonically with \( \mu \). Hence, the value \( \alpha_\ast \) given in \([19] \) is a lower bound, \( \alpha_\ast \leq \alpha < \infty \), while \( \beta_\ast \) quoted in \([20] \) is an upper bound, \( 0 < \beta \leq \beta_\ast \).

The above behavior extends to non-compact distributions with algebraic tails, \( \rho(x) \simeq b \mu x^{\mu - 1} \) when \( x \to \infty \). The condition \( \mu < -1 \) guarantees that the average is finite. In this case, the cumulative distribution has the tail \( R(x) \simeq b x^\mu \), and consequently, \( A_N \simeq (bN/\alpha)^{-1/\mu} \) with the \( \alpha \) given in \([22] \). Therefore, the exponent \( \beta \) shown in figure \([4] \) holds.

We also observe that the cumulative distributions \( F_N \) are polynomials of degree \( N \) in the function \( R(x) \). For \( N = 1 \), we have \( F_1 = R \) and hence \( S_1 = R_1 \). The first iteration of \([7] \) gives \( F_2 = R(1 + R_1 - R) \) and thus, \( S_2 = R_2(1 + R_1 - R_2) \). In general, the probabilities \( S_N \) are specified by the values \( R_n \) for \( n = 1, 2, \ldots, N \). For the uniform distribution \( R_N = 1/(N + 1) \); iteration of equation \([7] \) gives \( S_1 = 1/2 \), \( S_2 = 7/12 \), \( S_3 = 121/252 \) etc.

We now consider the dual probability \( I_N \) that all records are inferior, that is, they are below average: \( X_n < A_n \) for all \( n \leq N \). This quantity obeys the simple recursion
\[
I_{N+1} = I_N(1 - R_{N+1}), \tag{23}
\]
with \( I_1 = 1 - R_1 \). The factor \( 1 - R_{N+1} \) guarantees that the record \( X_{N+1} \) is inferior, regardless of the history of the sequence. In contrast with the recursion \([7] \), the probability \( I_N \) obeys a closed equation. The solution is the product
\[
I_N = (1 - R_1)(1 - R_2) \cdots (1 - R_N). \tag{24}
\]
For the uniform distribution, \( I_N = \frac{1}{2} \frac{2}{3} \cdots \frac{N}{N+1} \), an therefore, \( I_N = R_N = 1/(N + 1) \).

To obtain the asymptotic behavior for an arbitrary distribution, we convert the difference equation \([23] \) into the differential equation \( dI/dN = -\alpha I/N \). The probability \( I \) decays algebraically,
\[
I \sim N^{-\alpha}, \tag{25}
\]
with the exponent \( \alpha \) given by \([10] \). Indeed, for the uniform distribution, we recover \( \alpha = 1 \). Once again, the tail of the distribution \( \rho(x) \) controls the exponent \( \alpha \) (see also figure \([1] \)).

In summary, we studied statistics of superior records in a sequence of uncorrelated random variables. According to our definition, a sequence of records is superior if all records are above average. We presented a general theoretical framework that applies for arbitrary probability distribution functions, and used scaling methods to analyze the asymptotic behavior of large sequences. We obtained analytically the distribution of records and the fraction of superior sequences. The latter quantity decays algebraically with sequence length. Interestingly, the decay exponent is nontrivial, and it is controlled by the tail of the probability distribution function from which the random variables are drawn.

Our results show that first-passage properties of records are quite rich. Our study compares the actual record with the average expected for a given distribution as a probe of performance. Yet, performance is only one in a larger family of characteristics involving the entire history of the sequence. Our results suggest that there are additional “persistence”-like exponents \([22-23] \) for record sequences. Finally, it will be interesting to investigate superior records in sequences of correlated random variables, e.g. when the sequence \( x_n \) represents a random walk \([24] \).

We thank Joan Gomberg for useful discussions and the IAS (University of Warwick) for hospitality, and acknowledge DOE grant DE-AC52-06NA25396 for support.

[1] E. Castillo, Extreme Value Theory in Engineering (Academic Press, New York, 1988).
[2] P. Embrechts, G. Klüppelberg and T. Mikosch, Modelling extemeal events for insurance and finance (Spring-Verlag, Berlin, 1997).
[3] S. Y. Novak, Extreme value methods with applications to finance (Chapman & Hall/CRC Press, London, 2011).
[4] W. Feller, An Introduction to Probability Theory and Its Applications (Wiley, New York, 1968).
[5] R. S. Ellis, Entropy, Large Deviations, and Statistical Mechanics (Springer, Berlin 2005).
[6] E. I. Gumble, Statistics of Extremes (Dover, New York 2004).
[7] P. L. Krapivsky, S. Redner and E. Ben-Naim, A Kinetic View of Statistical Physics (Cambridge University Press, Cambridge, UK, 2010).
[8] R. A. Fisher and L. H. C. Tippett, Proc. Cambridge Phil. Soc. 24, 180 (1928).
[9] E. I. Gumble, Ann. Inst. Henri Poincaré 5, 115 (1935).
[10] S. Redner, A Guide to First-Passage Processes (Cambridge University Press, New York, 2001).
[11] A. J. Bray, S. N. Majumdar, and G. Schehr, Persistence and First-Passage Properties in Non-equilibrium Systems, arXiv:1304.1195.
[12] J. F. Eichner, E. Koscielny-Bunde, A. Bunde, S. Havlin, H. J. Schellnhuber, Phys. Rev. E 68, 046133 (2003).
[13] S. Redner and M. R. Petersen, Phys. Rev. E 74, 061114 (2006).
[14] W. I. Newman, B. D. Malamud, and D. L. Turcotte, Phys. Rev. E 82, 066111 (2010).
[15] J.-P. Bouchaud and M. Potters, Theory of Financial Risk and Derivative Pricing (Cambridge University Press, Cambridge 2003).
[16] Y. H. Liu, P. Gopikrishnan, P. Cizeau, M. Meyer, C. K. Peng, and H. E. Stanley, Phys. Rev. E 60, 1390 (1999).
[17] G. Wergen, M. Bogner, and J. Krug, Phys. Rev. E 83, 051109 (2011).
[18] J. Krug and K. Jain, Physica A 358, 1 (2005).
[19] N. G. van Kampen, Stochastic Processes in Physics and Chemistry (North-Holland, Amsterdam, 2001).
[20] R. L. Graham, D. E. Knuth, and O. Patashnik, Concrete Mathematics: A Foundation for Computer Science (Reading, Mass.: Addison-Wesley, 1989).
[21] E. Ben-Naim, P. L. Krapivsky, and S. Redner, Phys. Rev. E 50, 822 (1994).
[22] B. Derrida, V. Hakim, and V. Pasquier, Phys. Rev. Lett. 75, 751 (1995).
[23] S. N. Majumdar, C. Sire, A. J. Bray, and S. J. Cornell, Phys. Rev. Lett. 77, 2867 (1996).
[24] G. Wergen, S. N. Majumdar, and G. Schehr, Phys. Rev. E 86, 011119 (2012); G. Schehr and S. N. Majumdar, arXiv:1305.0639.