Probability in the Engineering and Informational Sciences, **19**, 2005, 289–308. Printed in the U.S.A.

EXTREMES

A CONTINUOUS-TIME PERSPECTIVE

IDDO ELIAZAR
Department of Mathematics
Bar-Ilan University
Ramat-Gan 52900, Israel
E-mail: eliazar@math.biu.ac.il; eliazar@post.tau.ac.il

We consider a generic continuous-time system in which events of random magnitudes occur stochastically and study the system's extreme-value statistics. An event is described by a pair (t, x) of coordinates, where t is the time at which the event took place and x is the magnitude of the event. The stochastic occurrence of the events is assumed to be governed by a Poisson point process.

We study various issues regarding the system's extreme-value statistics, including (i) the distribution of the largest-magnitude event, the distribution of the nth "runner-up" event, and the multidimensional distribution of the "top n" extreme events, (ii) the internal hierarchy of the extreme-value events—how large are their magnitudes when measured relative to each other, and (iii) the occurrence of record times and record values. Furthermore, we unveil a hidden Poissonian structure underlying the system's sequence of order statistics (the largest-magnitude event, the second largest event, etc.). This structure provides us with a markedly simple simulation algorithm for the entire sequence of order statistics.

1. INTRODUCTION

The Gaussian (Normal) curve is well known as the universal probability law governing the statistical distribution of large samples. However, what truly impacts us is not the overwhelming majority of "normal events," but the few rare exceptions of "abnormal" extreme events. Indeed, it is the unique Michael Jordan we remember, rather than a multitude of excellent professional NBA basketball players scoring an average number of points per game; it is hurricane Andrew that insurance compa-

 0269-9648/05 $16.00

https://doi.org/10.1017/S0269964805050163 Published online by Cambridge University Press

nies (painfully) recall, rather than the thousands of strong storms that took place in the United States in the 1990s; and it is the 1912 *Titanic* disaster we remember, rather than numerous other unfortunate sinking events of ships in the Atlantic. For many more examples of the importance we attribute to extreme events, we refer the readers to the *Guinness Book of Records*.

The interest in rare and extreme events was shared—apart from the devoted readers of the *Guinness Book of Records*—also by the scientific community. The pioneering studies took place in the 1920s and 1930s, with the works of von Bortkiewicz [20], Fréchet [8], Fisher and Tippett [7], von Mises [21], Weibull, and Gumbel [12]. A rigorous theoretical framework was presented in 1943 by Gnedenko [11]. Nowadays, the study of extremes is a well-established branch of Probability Theory called *Extreme Value Theory* (EVT). This theory is of major importance in the analysis of rare and "catastrophic" events such as floods in hydrology, large claims in insurance, crashes in finance, material failure in corrosion analysis, and so forth. Classic references on EVT are [9,12]. For both the theory and applications of modern EVT we refer the readers to [3,15,19].

Given a sequence $\{\xi_n\}_{n=1}^{\infty}$ of independent and identically distributed (i.i.d.) random variables (random samples), the "normal" approach is to study the asymptotic behavior of the scaled sums of the ξ's, namely the limiting probability distribution of

$$\hat{S}_n := \frac{\{\xi_1 + \cdots + \xi_n\} - b_n}{a_n} \tag{1}$$

(as $n \to \infty$), where $\{a_n\}_{n=1}^{\infty}$ and $\{b_n\}_{n=1}^{\infty}$ are properly chosen scaling coefficients. This, as the Central Limit Theorem asserts, leads to the universal Gaussian (Normal) distribution. (In case the ξ's fail to have a finite variance or mean, the Central Limit Theorem is replaced by the Generalized Central Limit Theorem, and the limiting distributions are the stable Lévy laws [13].) The "extreme" approach, on the other hand, studies the asymptotic behavior of the scaled *maxima* of the ξ's:

$$\hat{M}_n := \frac{\max\{\xi_1, \ldots, \xi_n\} - b_n}{a_n}. \tag{2}$$

The "Central Limit Theorem" of EVT asserts that (2) has three possible limiting probability laws (named, respectively, after the pioneers of EVT): Fréchet, Weibull, and Gumbel. The probability distribution functions of these universal extreme value laws are

$$\begin{aligned} &\textit{Fréchet:} \quad \exp\{-x^{-\alpha}\}, \qquad x > 0; \\ &\textit{Weibull:} \quad \exp\{-|x|^{\alpha}\}, \qquad x < 0; \\ &\textit{Gumbel:} \quad \exp\{-e^{-x}\}, \qquad x \text{ real} \end{aligned} \tag{3}$$

(the exponent α appearing in the Fréchet and Weibull distributions is a positive parameter).

https://doi.org/10.1017/S0269964805050163 Published online by Cambridge University Press

Typically, EVT studies the extreme statistics (maxima, minima, order statistics, records, etc.) of random sequences $\{\xi_n\}_{n=1}^{\infty}$. The fundamental theory considers i.i.d. sequences, but generalizations to stationary sequences do exist (see, for example, [3]). Such random sequences represent a *time series* of data measured at *discrete*-time epochs. However, most "real-world" systems are usually *continuous-time* systems. Hence, why not study the extremes of continuous-time systems *directly* in some appropriate continuous-time setting?

Consider a generic continuous-time physical system in which events that take place are monitored and logged. Each event is described by a pair (t, x) of coordinates, with t being the *time* at which the event occurred and x being the *magnitude* of the event (a numerical value). Hence, the "history" of the physical system is given by a random collection $\mathcal{X}$ of points in the plane, where each point of $\mathcal{X}$ represents an event that took place. In the mathematical nomenclature, the random collection $\mathcal{X}$—the system's history of events—is called a *point process*. A direct continuous-time approach would thus be to study the statistics of the extreme points of $\mathcal{X}$. To that end, one obviously has to specify the probability distribution of the point process $\mathcal{X}$. The continuous-time counterparts of discrete-time i.i.d. sequences are *Poisson point processes*. If $\mathcal{X}$ is a Poisson point process governed by the *rate function* $\lambda(x)$, then, informally, (i) events of magnitude belonging to the infinitesimal range $(x, x + dx)$ arrive at rate $\lambda(x)dx$ and (ii) the occurrences of events of different magnitudes are independent.

In this article, we will explore the extremes of a generic continuous-time physical system whose history of events forms a Poisson point process. As we will demonstrate, this setting turns out to be "tailor-made" to the modeling and analysis of extreme events in continuous time. We begin, in Section 2, with a short review of the notion of Poisson point processes and introduce our underlying continuous-time system model. In Section 3, we define the system's sequence of *order statistics* and explore their distributions: the probability law of the maximum, the probability law of the nth "runner-up," and the multidimensional probability law of the "top n" extremes. In Section 4, we unveil a hidden *Poissonian structure* underlying the sequence of order statistics. This hidden structure gives rise to a markedly simple *simulation algorithm* for the sequence of order statistics. In Section 5, we turn to study the *internal hierarchy* of the sequence of order statistics; namely we analyze the magnitudes of the extreme events when measured *relative* to each other. We conclude, in Section 6, with the exploration of the system's *records times* and *record values*.

A note about notation: Throughout the article, $\mathbb{R}$ denotes the real line, $\mathbf{P}(\cdot) :=$ probability, $\mathbf{E}[\cdot] :=$ expectation, and $d\omega$ (dx, dt, etc.) is used to denote the infinitesimal neighborhood of the point ω (x, t, etc.).

2. THE CONTINUOUS-TIME SETTING

In this section we concisely review the notion of Poisson point processes, introduce our underlying continuous-time "*event process*" (which will accompany us throughout the manuscript), and explain the passage from the discrete-time i.i.d. setting to the continuous-time Poisson setting.

2.1. Poisson Point Processes

Let Ω be an Euclidean space (or subspace or domain). A random countable collection of points $\Pi \subset \Omega$ is called a *point process* [17]. We denote by $\Pi(B)$ the number of points of Π residing in the subset B:

$$\Pi(B) = \text{card}\{\Pi \cap B\}. \tag{4}$$

Hence, the point process $\Pi \subset \Omega$ induces a *counting measure* on Ω given by (4).

A point process $\Pi \subset \Omega$ is said to be a *Poisson point process* [17] with rate $r(\omega)$ if the following pair of conditions hold:

- If B is a subset of Ω, then the random variable $\Pi(B)$ is *Poisson distributed* with mean $\int_B r(\omega)\, d\omega$.
- If $\{B_k\}_k$ is a finite collection of *disjoint* subsets of Ω, then $\{\Pi(B_k)\}_k$ is a finite collection of *independent* random variables.

The Poisson point process Π can also be described by its finite-dimensional distributions: The probability that a given set of points $\{\omega_j\}_{j=1}^n$ belongs to Π is

$$\mathbf{P}(\Pi(d\omega_1) = 1, \ldots, \Pi(d\omega_n) = 1) = r(\omega_1)\, d\omega_1 \ldots r(\omega_n)\, d\omega_n. \tag{5}$$

Informally, Ω is divided into infinitesimal cells. Each cell contains either a single point or no points at all. The cells are independent, and the probability that the cell $d\omega$ contains a point is $r(\omega)\, d\omega$.

The best known example of a Poisson point process is the *standard Poisson process*, where $\Omega = [0,\infty)$ and the rate function $r(\omega)$ is constant.

2.2. The Event Process

Equipped with the notion of Poisson point processes, we can now rigorously define the point process $\mathcal{X}$ presented in Section 1. Recall that we considered a generic continuous-time physical system whose "history" is a random collection $\mathcal{X}$ of points in the plane—the point (t, x) represents an event of magnitude x taking place at time t. In Section 1, we said, informally, that "events of magnitude belonging to the infinitesimal range $(x, x + dx)$ arrive at rate $\lambda(x)\, dx$" and, "the occurrences of events of different magnitudes are independent." Put rigorously, the process $\mathcal{X}$ (henceforth referred to as our *event process*) is taken to be a time-homogeneous Poisson point process on $\Omega = [0,\infty) \times \mathbb{R}$ with rate

$$r(t, x) = \lambda(x). \tag{6}$$

Thus,

$$(t, x) \in \mathcal{X} \quad \text{with probability } \lambda(x)\, dt\, dx. \tag{7}$$

https://doi.org/10.1017/S0269964805050163 Published online by Cambridge University Press

We set $\Lambda(x)$ to be the Poissonian rate at which samples of size *greater* than x arrive, namely

$$\Lambda(x) := \int_x^\infty \lambda(u)\,du. \tag{8}$$

The function $\Lambda(x)$ is smooth, monotone nonincreasing, and fully characterizes the event process $\mathcal{X}$. This function will turn out to play a key role in the sequel. We henceforth refer to the function $\Lambda(x)$ as the *characteristic* of the process $\mathcal{X}$ and assume that

$$\lim_{x\to-\infty} \Lambda(x) = +\infty \quad \text{and} \quad \lim_{x\to+\infty} \Lambda(x) = 0. \tag{9}$$

We will elaborate on this assumption in the following subsection. Furthermore, in Subsection 3.1 we shall show that this assumption is in fact an essential requirement.

Finally, note that the random variable $\mathcal{X}([t,t+\Delta]\times[a,b])$—counting the number of events of size $x\in[a,b]$ occurring during the time interval $[t,t+\Delta]$—is Poisson distributed with mean

$$\int_t^{t+\Delta}\int_a^b r(t,x)\,dt\,dx = \Delta\int_a^b \lambda(x)\,dx = \Delta(\Lambda(a)-\Lambda(b)). \tag{10}$$

2.3. From Discrete to Continuous and Back

In the discrete-time setting described in Section 1, the underlying time series $\{\xi_n\}_{n=1}^\infty$ is an i.i.d. sequence of random variables. If we take the random samples $\{\xi_n\}_{n=1}^\infty$ to arrive according to some continuous-time counting process $(N(t))_{t\ge0}$, then the sample set, at time t, would be

$$\{\xi_1,\ldots,\xi_{N(t)}\} \tag{11}$$

(the sample set being empty in the case $N(t)=0$). The setting in which $(N(t))_{t\ge0}$ is a *renewal process* was introduced in [10] and coined "Random Record Models" (see also [4,22]).

If $(N(t))_{t\ge0}$ is a standard Poisson process with rate ρ, then the "sample process" (11) is in fact a time-homogeneous Poisson point process on $[0,\infty)\times\mathbb{R}$ with rate

$$r(t,x) = \rho f(x),$$

where $f(x)$ is the probability density function of the ξ's.

On the other hand, if the rate function $\lambda(x)$ of the event process $\mathcal{X}$ is integrable (i.e., if $\int\lambda(x)\,dx<\infty$), then $\mathcal{X}$ can be described in the form of the sample set (11), where (i) the Poissonian arrival rate is $\rho := \int\lambda(x)\,dx$ and (ii) the probability density function of the ξ's is $f(x) := \lambda(x)/\rho$.

Hence, in the case of integrable rate functions, there is a one-to-one correspondence, via embedding, between discrete-time i.i.d. sequences $\{\xi_n\}_{n=1}^\infty$ and continuous-time event processes $\mathcal{X}$. However, the truly interesting case is where the rate function is actually *not* integrable: $\int \lambda(x)\,dx = \infty$. In this case, $\mathcal{X}$ has *infinitely many* events occurring on *all time scales* (i.e., on any time interval), and *no* correspondence between $\mathcal{X}$ and any discrete-time i.i.d. sequence $\{\xi_n\}_{n=1}^{\infty}$ can be established. Thus, in the case of a nonintegrable rate function, the event process $\mathcal{X}$ is a truly continuous-time "creature."

In this article, we focus on the "truly continuous-time case," where the rate functions are *nonintegrable*. This is the reason for assumption (9). The limit $\lim_{x\to-\infty}\Lambda(x) = +\infty$ ensures that the rate function $\lambda(x)$ is indeed nonintegrable. The limit $\lim_{x\to+\infty}\Lambda(x) = 0$, on the other hand, ensures that events of infinitely large magnitude do not occur. We will return to discuss assumption (9) in Subsection 3.1.

3. THE SEQUENCE OF ORDER STATISTICS

Let

$$X_1(t) > X_2(t) > X_3(t) > \cdots \tag{12}$$

denote the sequence of *order statistics* of the event process $\mathcal{X}$; namely $X_n(t)$ is the nth largest sample observed during the time period $[0,t]$:

$$X_n(t) = \sup\{x < X_{n-1}(t) \,|\, (s,x) \in \mathcal{X} \text{ and } s \le t\}, \tag{13}$$

where $X_0(t)$ is set to equal $+\infty$. (Definition (13) is in fact valid for any point process on $\Omega = [0,\infty) \times \mathbb{R}$.)

Let us now turn to analyze the probability distributions of the maximum value $X_1(t)$, the nth "runner-up" $X_{n+1}(t)$, and the vector of the "top n" order statistics $(X_1(t),\ldots,X_n(t))$.

3.1. The Maximum

We compute the distribution of the maximum $X_1(t)$. The maximum at time t is less or equal to x if and only if no events of magnitude greater than x occurred during the time period $[0,t]$; that is,

$$\{X_1(t) \le x\} = \{\mathcal{X}([0,t] \times (x,\infty)) = 0\}.$$

However, the random variable $\mathcal{X}([0,t] \times (x,\infty))$ is Poisson distributed with mean $t\Lambda(x)$ (recall (10)) and, hence,

$$\mathbf{P}(X_1(t) \le x) = \exp\{-t\Lambda(x)\}. \tag{14}$$

The probability density function of $X_1(t)$ is given, in turn, by

$$\exp\{-t\Lambda(x)\}t\lambda(x). \tag{15}$$

The assumption of (9): From (14), we see that the distribution of the maximum is proper if and only if the assumption of (9) holds. Indeed, in order for the distri-

https://doi.org/10.1017/S0269964805050163 Published online by Cambridge University Press

bution of $X_1(t)$ to be proper, the limit of (14) must equal zero when $x \to -\infty$ and must equal one when $x \to +\infty$. This, however, takes place if and only if $\lim_{x\to-\infty}\Lambda(x) = +\infty$ and $\lim_{x\to+\infty}\Lambda(x) = 0$. Hence, the assumption of (9) is not an impeding restriction, but an essential requirement.

3.2. Beating the Maximum

Assume that the current maximum level is x. How long will we have to wait until this maximum level is beat? Since events of magnitude greater than x arrive at rate $\Lambda(x)$, the answer is straightforward: The waiting time is exponentially distributed with rate $\Lambda(x)$ (mean $1/\Lambda(x)$). Let us consider the two following variations of this question.

(i) Assume that we are at time t, but we *do not* know the current maximum level $X_1(t)$. How long will we have to wait—from the present time t onward—until the unknown maximum level $X_1(t)$ is beat? Let $L(t)$ denote the respective waiting time. Given that $X_1(t) = x$, the waiting time $L(t)$ is exponentially distributed with rate $\Lambda(x)$. Conditioning on the (unknown) maximum level $X_1(t)$, we obtain that

$$\mathbf{P}(L(t) > l) = \frac{1}{1 + l/t}. \tag{16}$$

The proof of (16) is given below. Note that the waiting time $L(t)$ has *infinite mean*.

(ii) This variation regards positive-valued event processes. Assume (as earlier) that the current maximum level is x and let $k > 1$ be a fixed parameter. How long will we have to wait until the occurrence of a maximum level that is at least k times larger than *all* the maximum levels preceding it? These waiting times (coined "geometric record times") are explored in [2] (their analysis is considerably harder than the analysis of the waiting times described earlier).

To prove (16), use the probability density function of (15) and the change of variables $u = t\Lambda(x)$:

$$\begin{aligned}\mathbf{P}(L(t) > l) &= \int_0^\infty \mathbf{P}(L(t) > l \,|\, X_1(t) = x)\mathbf{P}(X_1(t) \in dx)\\ &= \int_0^\infty \exp\{-l\Lambda(x)\}\exp\{-t\Lambda(x)\}t\lambda(x)\,dx\\ &= \int_0^\infty \exp\{-(1 + l/t)u\}\,du\\ &= (l + l/t)^{-1}.\end{aligned}$$

https://doi.org/10.1017/S0269964805050163 Published online by Cambridge University Press

3.3. Fréchet, Weibull, and Gumbel

Equation (14) implies that the function $\Lambda(x)$ fully characterizes the distribution of the maximum. On the other hand, the function $\Lambda(x)$ also fully characterizes the underlying event process $\mathcal{X}$. Hence, there is a *one-to-one correspondence*—conveyed by the characteristic $\Lambda(x)$—between event processes and the distribution of their maxima. In particular, the event processes corresponding to the "Central Limit Theorem" distributions of EVT (viz. the Fréchet, Weibull, and Gumbel distributions) are given, respectively, by

$$\begin{array}{llll} \textit{Fréchet}: & \Lambda(x) = x^{-\alpha}, & \lambda(x) = \alpha x^{-\alpha-1}, & x > 0; \\ \textit{Weibull}: & \Lambda(x) = |x|^{\alpha}, & \lambda(x) = \alpha |x|^{\alpha-1}, & x < 0; \\ \textit{Gumbel}: & \Lambda(x) = \exp\{-x\}, & \lambda(x) = \exp\{-x\}, & x \text{ real} \end{array}$$

(the exponent α appearing in the Fréchet and Weibull distributions is a positive parameter).

Could these extreme value laws be obtained (as in the "classic" EVT) as the only possible maxima scaling limits? The answer, as explained below, is affirmative.

The continuous-time scaling of the maximum $X_1(t)$, analogous to the discrete-time scaling of (2), is

$$\hat{X}_1(t) := \frac{X_1(t) - b(t)}{a(t)},$$

where $a(t)$ and $b(t)$ are the scaling functions ($a(t)$ being positive valued). The distribution of the scaled maxima $\hat{X}_1(t)$, using (14), is hence given by

$$\mathbf{P}(\hat{X}_1(t) \le x) = \exp\{-t\Lambda(a(t)x + b(t))\}. \tag{17}$$

On the other hand, the distribution of the scaled maxima $\hat{M}_n$ of (2) is given by

$$\mathbf{P}(\hat{M}_n \le x) = \exp\{-nL(a_n x + b_n)\}, \tag{18}$$

where $L(x) := -\ln(\mathbf{P}(\xi_1 \le x))$. Moreover, the function $L(x)$ satisfies the very same properties the function $\Lambda(x)$ does; namely it is monotone decreasing, with $\lim_{x\to-\infty} L(x) = +\infty$ and $\lim_{x\to+\infty} L(x) = 0$.

Thus, (17) (as $t \to \infty$) and (18) (as $n \to \infty$) must yield the *same* distributional limits. The "classic" EVT asserts that the three possible limits of (18) are the extreme value distributions: Fréchet, Weibull, and Gumbel. Hence, these probability laws are also the only possible limits of (17).

The scaling functions in the Fréchet, Weibull, and Gumbel cases are

$$\begin{array}{lll} \textit{Fréchet}: & a(t) = t^{1/\alpha}, & b(t) = 0; \\ \textit{Weibull}: & a(t) = t^{-1/\alpha}, & b(t) = 0; \\ \textit{Gumbel}: & a(t) = 1, & b(t) = \ln(t). \end{array}$$

https://doi.org/10.1017/S0269964805050163 Published online by Cambridge University Press

It is interesting to note that in these three special cases, the scaling yields $\mathbf{P}(\hat{X}_1(t) \leq x) = \exp\{-\Lambda(x)\}$ ($\forall t \geq 0$).

For a detailed analysis regarding the "basins of attraction" of the Fréchet, Weibull, and Gumbel extreme value distributions, we refer the readers to [3].

3.4. The *n*th "Runner-up" and the "Top *n*"

The distribution of $X_{n+1}(t)$—the nth "runner-up" is computed analogously to the distribution of the maximum $X_1(t)$. The $(n+1)$st order statistic at time t is less or equal to x if and only if no more than n events of magnitude greater than x occurred during the time period $[0,t]$; that is,

$$\{X_{n+1}(t) \leq x\} = \{\mathcal{X}([0,t] \times (x,\infty)) \leq n\}.$$

Hence, since $\mathcal{X}([0,t] \times (x,\infty))$ is Poisson distributed with mean $t\Lambda(x)$, we have

$$\mathbf{P}(X_{n+1}(t) \leq x) = \sum_{k=0}^{n} \frac{(t\Lambda(x))^k}{k!} \exp\{-t\Lambda(x)\}. \tag{19}$$

The probability density function of $X_{n+1}(t)$ is given, in turn, by

$$\frac{(t\Lambda(x))^n}{n!} \exp\{-t\Lambda(x)\} t\lambda(x). \tag{20}$$

Note that (19) and (20) indeed coincide with (14) and (15) in the case $n = 0$.

Equation (19) gives the *one-dimensional* distribution of the order statistics. What about the joint, *multidimensional*, probability distribution of the order statistics; namely the joint distribution of the "top n"—the vector $(X_1(t), \dots, X_n(t))$?

Well, the joint probability density function of the vector $(X_1(t), \cdots, X_n(t))$ is given by

$$f_n(x_1, \dots, x_n) = t^n \lambda(x_1) \cdots \lambda(x_n) \exp\{-t\Lambda(x_n)\}, \tag{21}$$

where $x_1 > x_2 > \cdots > x_n$. The explanation follows.

In order for the points $x_1 > x_2 > \cdots > x_n$ to be the "top n" points of the sample process $\mathcal{X}$ at time t, we need (for $j = 1, \dots, n$) the following: (i) an event of magnitude x_j takes place during the time period $[0,t]$; this occurs with probability $t\lambda(x_j)\,dx_j$; and (ii) no events of magnitude $\in (x_j, x_{j-1})$ (where x_0 is set to equal $+\infty$) take place during the time period $[0,t]$; this occurs with probability $\exp\{-t(\Lambda(x_j) - \Lambda(x_{j-1}))\}$. Multiplying these probabilities together yields the multidimensional density function (21).

4. THE STRUCTURE OF THE SEQUENCE OF ORDER STATISTICS

When viewed as a stochastic process in the "order parameter" n (keeping the time t fixed), the sequence of order statistics $\{X_n(t)\}_{n=1}^{\infty}$ conceals a hidden underlying structure, which we unveil in this section. First, we reveal a *Markovian* structure

https://doi.org/10.1017/S0269964805050163 Published online by Cambridge University Press

governing the sequence of order statistics. Second, we prove that the Markovian structure is due to a hidden, underlying, *Poissonian* structure. This Poissonian structure, in turn, provides us with a markedly simple *simulation algorithm* for the sequence of order statistics. Third, we show that the (discretely parameterized) sequence of order statistics can be embedded in a simple transformation of a (continuously parameterized) *Gamma process*.

4.1. Markovian Structure

The *conditional distribution* of a $(n+1)$st order statistic, *given* the value of the nth order statistic, is

$$\mathbf{P}(X_{n+1}(t) \le y \,|\, X_n(t) = x) = \exp\{-t(\Lambda(y) - \Lambda(x))\} \tag{22}$$

$(y < x)$. The explanation follows.

Given that the nth order statistic equals x, the $(n+1)$st order statistic will be less or equal to y (where $y < x$) if and only if no events of magnitude $\in (y, x)$ occur during the time interval $[0, t]$; that is, if and only if $\mathcal{X}([0,t] \times (y,x)) = 0$. Since $\mathcal{X}$ is a Poisson point process, the random variable $\mathcal{X}([0,t] \times (y,x))$ is Poisson distributed with mean $t(\Lambda(y) - \Lambda(x))$ and is *independent* of the points of $\mathcal{X}$ residing in $[0,t] \times [x,\infty)$. Hence, the left-hand side of (22) equals the probability that $\mathcal{X}([0,t] \times (y,x)) = 0$, which, in turn, is given by the right-hand side of (22).

Since (22) is a *Markovian* recursion, it implies that the sequence of order statistics $\{X_n(t)\}_{n=1}^{\infty}$ is a *Markov chain* (in the variable n, for t fixed). The initial condition of this Markov chain is given by the distribution of the maximum $X_1(t)$; see (14).

4.2. The Hidden Poissonian Structure

When viewed in the proper perspective, the sequence of order statistics $\{X_n(t)\}_{n=1}^{\infty}$ conceals a hidden *Poissonian* structure. The "proper perspective" is to consider the sequence $\{\Lambda(X_n(t))\}_{n=1}^{\infty}$, rather than the original sequence $\{X_n(t)\}_{n=1}^{\infty}$. Let us begin with the distribution of $\Lambda(X_1(t))$.

Using (14), we have

$$\mathbf{P}(\Lambda(X_1(t)) > u) = \mathbf{P}(X_1(t) < \Lambda^{-1}(u)) = \exp\{-t\Lambda(\Lambda^{-1}(u))\} = \exp\{-tu\}$$

$(u \ge 0)$. Hence, $\Lambda(X_1(t))$ is exponentially distributed with parameter t (mean $1/t$). In a similar way, (19) implies that $\Lambda(X_n(t))$ is Gamma distributed with parameters (t, n).

More informative, however, is to analyze the *conditional distribution* of the increment $\Lambda(X_{n+1}(t)) - \Lambda(X_n(t))$ *given* $\Lambda(X_n(t))$. Indeed, using (22), we have

$$\begin{aligned}
&\mathbf{P}(\Lambda(X_{n+1}(t)) - \Lambda(X_n(t)) > u \,|\, \Lambda(X_n(t)) = v)\\
&\quad = \mathbf{P}(X_{n+1}(t) < \Lambda^{-1}(u+v) \,|\, X_n(t) = \Lambda^{-1}(v))\\
&\quad = \exp\{-t(\Lambda(\Lambda^{-1}(u+v)) - \Lambda(\Lambda^{-1}(v)))\}\\
&\quad = \exp\{-tu\};
\end{aligned}$$

https://doi.org/10.1017/S0269964805050163 Published online by Cambridge University Press

that is, the increment $\Lambda(X_{n+1}(t)) - \Lambda(X_n(t))$ is *independent* of $\Lambda(X_n(t))$ and is exponentially distributed with parameter t (mean $1/t$). Hence, we have obtained the following proposition:

PROPOSITION 1: *Let $\mathcal{X}$ be an event process with characteristic $\Lambda(x)$ and order statistics $\{X_n(t)\}_{n=1}^{\infty}$. Then the increasing sequence of points $\{\Lambda(X_n(t))\}_{n=1}^{\infty}$ forms a standard Poisson process with rate t.*

4.3. Simulation and Gamma Embedding

Proposition 1 provides us with a remarkably simple simulation algorithm for the *entire sequence* of order statistics $\{X_n(t)\}_{n=1}^{\infty}$:

$$X_n(t) = \Lambda^{-1}\left(\frac{Z_1 + Z_2 + \cdots + Z_n}{t}\right), \tag{23}$$

where $\{Z_n\}_{n=1}^{\infty}$ is an i.i.d. sequence of exponentially distributed random variables with unit mean. In particular, the simulation algorithms for the Fréchet, Weibull, and Gumbel order statistics are as follows:

$$\textit{Fréchet}: \quad X_n(t) = \left(\frac{t}{Z_1 + Z_2 + \cdots + Z_n}\right)^{1/\alpha};$$

$$\textit{Weibull}: \quad X_n(t) = -\left(\frac{Z_1 + Z_2 + \cdots + Z_n}{t}\right)^{1/\alpha};$$

$$\textit{Gumbel}: \quad X_n(t) = \ln(t) - \ln(Z_1 + Z_2 + \cdots + Z_n).$$

Equation (23) is in fact a manifestation of a *Gamma embedding* as we now explain.

The *Gamma process* $(G(s))_{s\geq 0}$ is a stochastic process starting at the origin ($G(0) = 0$) whose increments are *independent*, *stationary*, and *Gamma distributed*:

$$\mathbf{P}(G(b) - G(a) \in dx) = \frac{1}{\Gamma(b-a)} \exp\{-x\} x^{(b-a)-1}\, dx$$

($b > a \geq 0$). The Gamma process is a special example of one-sided Lévy processes (Lévy subordinators) [1].

At the point $s = 1$, the value of the Gamma process $G(1)$ is exponentially distributed with unit mean. Hence, since the Gamma process has independent and stationary increments, the increments $\{G(n) - G(n-1)\}_{n=1}^{\infty}$ form an i.i.d. sequence of exponentially distributed random variables with unit mean. Combining this observation with (23), we obtain the following Gamma embedding of the sequence of order statistics $\{X_n(t)\}_{n=1}^{\infty}$:

$$X_n(t) = \Lambda^{-1}\left(\frac{G(n)}{t}\right); \tag{24}$$

that is, the discrete-parameter sequence $\{X_n(t)\}_{n=1}^{\infty}$ is embedded in the continuous-time process $(\Lambda^{-1}(G(s)/t))_{s\geq 0}$, which, in turn, is a transformation of the Gamma process $(G(s))_{s\geq 0}$.

Furthermore, the random variable $\Lambda^{-1}(G(s)/t)$, for noninteger parameter values s, may be considered a "virtual" s-order statistic. Continuing yet another step in this direction, we can define the Fréchet, Weibull, and Gumbel *virtual order processes* as follows ($s \geq 0$):

$$\begin{aligned} &\textit{Fréchet}: && G(s)^{-1/\alpha}; \\ &\textit{Weibull}: && -G(s)^{1/\alpha}; \\ &\textit{Gumbel}: && -\ln(G(s)). \end{aligned}$$

5. THE INTERNAL HIERARCHY OF THE SEQUENCE OF ORDER STATISTICS

What is the "internal hierarchy" of the sequence of the order statistics? What are the relative magnitudes of the order statistics? The answer to these questions is given by the following proposition.

PROPOSITION 2: *Let $\mathcal{X}$ be an event process with characteristic $\Lambda(x)$ and order statistics $\{X_n(t)\}_{n=1}^{\infty}$. Then the ratios $\{\Lambda(X_n(t))^n/\Lambda(X_{n+1}(t))^n\}_{n=1}^{\infty}$ are independent and uniformly distributed on the unit interval. Equivalently,*

$$\mathbf{P}\left(\frac{\Lambda(X_n(t))}{\Lambda(X_{n+1}(t))} \leq u\right) = u^n \qquad (0 < u < 1). \tag{25}$$

The proof of Proposition 2, which is based on the use of an order-preserving transformation of event processes, is provided in the Appendix.

We note that the *distribution* of the ratio $\Lambda(X_n(t))/\Lambda(X_{n+1}(t))$ can be deduced from Proposition 1. Indeed, (23) implies that this ratio equals, in law, the ratio $(Z_1 + \cdots + Z_n)/(Z_1 + \cdots + Z_{n+1})$, where $\{Z_n\}_{n=1}^{\infty}$ is an i.i.d. sequence of exponentially distributed random variables with unit mean. The latter ratio, in turn, is known (see, for example, [6]) to be governed by the Beta distribution function u^n $(0 < u < 1)$.

With Proposition 2 at hand, let us explore the internal hierarchy of the sequence of order statistics in the special Fréchet, Weibull, and Gumbel cases.

5.1. The Fréchet and Weibull Cases

In the Fréchet case, $\Lambda(x) = x^{-\alpha}$, and in the Weibull case, $\Lambda(x) = |x|^{\alpha}$ (the exponent α being a positive parameter). Hence, Proposition 2 implies the following:

- **Fréchet:** The ratios $X_{n+1}(t)/X_n(t)$, $n = 1,2,\ldots$, are *independent* random variables governed by the Beta distribution

$$\mathbf{P}\left(\frac{X_{n+1}(t)}{X_n(t)} \leq u\right) = u^{\alpha n}. \tag{26}$$

- **Weibull:** The ratios $|X_n(t)|/|X_{n+1}(t)|$, $n = 1,2,\ldots$, are *independent* random variables governed by the Beta distribution

$$\mathbf{P}\left(\frac{|X_n(t)|}{|X_{n+1}(t)|} \le u\right) = u^{\alpha n}. \tag{27}$$

Most interesting is the first ratio—$X_2(t)/X_1(t)$ in the Fréchet case and $|X_1(t)|/|X_2(t)|$ in the Weibull case—which measures the relative magnitudes of the "winner" and the "first runner-up." The ratio distribution, in both the Fréchet and Weibull cases, is governed by the probability density function $f(u) = \alpha u^{\alpha-1}$ $(0 < u < 1)$. This density function undergoes a phase transition when crossing the parameter value $\alpha = 1$: (i) When $0 < \alpha < 1$, the density $f(u)$ is monotone decreasing, starting at $f(0) = \infty$ and decreasing to $f(1) = \alpha$; (ii) when $\alpha = 1$ (the "critical" parameter value), the density $f(u)$ is uniform; and (iii) when $\alpha > 1$, the density $f(u)$ is monotone increasing, starting at $f(0) = 0$ and increasing to $f(1) = \alpha$. Hence, the "first runner-up" trails considerably behind the "winner" when α is small, and it follows the "winner" closely when α is large.

At the other end of the order spectrum, the distribution function $u^{\alpha n}$ becomes more and more concentrated around the value $u = 1$ as $n \to \infty$. This implies that as $n \to \infty$, the order statistics get closer and closer to each other *in ratio*. In the Fréchet case, as $n \to \infty$, the order statistics converge to zero and hence get closer and closer to each other, also absolutely. In the Weibull case, on the other hand, the order statistics diverge to $-\infty$ and hence remain apart.

In the Fréchet and Weibull cases, Proposition 2 also provides us with an algorithm for the simulation of the *relative* contributions of the "top n" order statistics $X_1(t), X_2(t), \ldots, X_n(t)$ to their *aggregate* magnitude. The explanation follows.

Let $\{U_k\}_{k=1}^{n-1}$ be an i.i.d. sequence of uniformly distributed random variables, and set $\{M_k\}_{k=0}^{n-1}$ to be given by

$$M_k = \begin{cases} 1, & k = 0 \\ U_1^{1/\alpha} U_2^{1/2\alpha} \cdots U_k^{1/k\alpha}, & k > 0. \end{cases}$$

Proposition 2 implies (all equalities below are equalities in law):

Fréchet case: $X_{k+1}(t)/X_k(t) = U_k^{1/k\alpha}$ and hence, using recursion, $X_k(t) = X_1(t)M_{k-1}$. This, in turn, yields the arithmetic average:

$$\frac{(X_1(t), X_2(t), \ldots, X_n(t))}{X_1(t) + X_2(t) + \cdots + X_n(t)} = \frac{(M_0, M_1, \ldots, M_{n-1})}{M_0 + M_1 + \cdots + M_{n-1}}. \tag{28}$$

Weibull case: $|X_k(t)|/|X_{k+1}(t)| = U_k^{1/k\alpha}$ and hence, using recursion, $|X_k(t)| = |X_n(t)|M_{n-1}/M_{k-1}$. This, in turn, yields the harmonic average:

$$\frac{(|X_1(t)|, |X_2(t)|, \ldots, |X_n(t)|)}{|X_1(t)| + |X_2(t)| + \cdots + |X_n(t)|} = \frac{(M_0^{-1}, M_1^{-1}, \ldots, M_{n-1}^{-1})}{M_0^{-1} + M_1^{-1} + \cdots + M_{n-1}^{-1}}. \tag{29}$$

5.2. The Gumbel Case

In the Gumbel case $\Lambda(x) = \exp\{-x\}$ and hence proposition 2 implies:

> **Gumbel:** The differences $X_n(t) - X_{n+1}(t)$, $n = 1,2,\ldots$, are *independent* random variables governed by the exponential distribution
>
> $$\mathbf{P}(X_n(t) - X_{n+1}(t) > u) = \exp\{-nu\}. \tag{30}$$

Alternatively, the random variables $\{n(X_n(t) - X_{n+1}(t))\}_{n=1}^{\infty}$ are i.i.d. and exponentially distributed with unit mean.

Let us study the following "asymptotically centered" sequence of differences $\{D_n\}_{n=1}^{\infty}$ defined by

$$D_n := X_1(t) - X_{n+1}(t) - \ln(n).$$

Using (30), the sequence $\{D_n\}_{n=1}^{\infty}$ admits the probabilistic representation

$$D_n = \frac{Z_1}{1} + \frac{Z_2}{2} + \cdots + \frac{Z_n}{n} - \ln(n), \tag{31}$$

where $\{Z_n\}_{n=1}^{\infty}$ is an i.i.d. sequence of exponentially distributed random variables with unit mean. The representation (31), in turn, yields that the mean and variance of D_n are

$$\mathbf{E}[D_n] = \frac{1}{1} + \frac{1}{2} + \cdots + \frac{1}{n} - \ln(n) \xrightarrow[n\to\infty]{} \gamma,$$

$$\mathrm{Var}[D_n] = \frac{1}{1^2} + \frac{1}{2^2} + \cdots + \frac{1}{n^2} \xrightarrow[n\to\infty]{} \frac{\pi^2}{6},$$

where $\gamma \simeq 0.577$ is Euler's constant.

Moreover, the representation (31) further yields that the Laplace transform of D_n is

$$\mathbf{E}[\exp\{-\theta D_n\}] = \frac{1}{n^\theta}\,\frac{1}{1+\theta}\,\frac{2}{2+\theta}\cdots\frac{n}{n+\theta} = \frac{\Gamma(1+\theta)}{n^\theta}\,\frac{\Gamma(n+1)}{\Gamma(n+1+\theta)}$$

($\theta \geq 0$). Hence, using Stirling's formula, we have

$$\mathbf{E}[\exp\{-\theta D_n\}] \xrightarrow[n\to\infty]{} \Gamma(1+\theta).$$

The function $\Gamma(1+\theta)$, however, is the Laplace transform of the standard Gumbel distribution (see, for example, [5]), and hence we conclude the following:

The sequence $\{D_n\}_{n=1}^{\infty}$ converges, in law, to the standard Gumbel distribution.

6. RECORDS

Among the points of the event process $\mathcal{X}$, the set of *record points* (henceforth denoted by $\mathcal{R}$) is of special interest. This last section is devoted to the study of these points.

A point $(t,x) \in \mathcal{X}$ is said to be a record point if and only if *all* events occurring during the time period $[0,t)$ were *smaller* in magnitude than the value x:

$$(t,x) \in \mathcal{R} \Leftrightarrow \mathcal{X}([0,t) \times [x,\infty)) = 0. \tag{32}$$

If $(t,x) \in \mathcal{R}$, then t is called a *record time* and x is called a *record value*. We denote the sets of record times and record values respectively by $\mathcal{R}_{\text{time}}$ and $\mathcal{R}_{\text{value}}$. Clearly, the sets $\mathcal{R}_{\text{time}} \subset [0,\infty)$ and $\mathcal{R}_{\text{value}} \subset \mathbb{R}$ are the projections of the record set $\mathcal{R}$ on the time and space axes.

The probability that the points $\{(t_j,x_j)\}_{j=1}^n$, where $t_1 < t_2 < \cdots < t_n$ and $x_1 < x_2 < \cdots < x_n$, belong to the record set $\mathcal{R}$ is given by

$$\mathbf{P}(\mathcal{R}(dt_j \times dx_j) = 1\,; j = 1,\ldots,n) = \prod_{j=1}^n \exp\{-(t_j - t_{j-1})\Lambda(x_j)\}\lambda(x_j)\,dt_j\,dx_j, \tag{33}$$

where t_0 is set to equal zero. The explanation follows.

In order for the points $\{(t_j,x_j)\}_{j=1}^n$ to belong to the record set $\mathcal{R}$, we need the following (for $j = 1,\ldots,n$): (i) the point (t_j,x_j) belongs to the underlying event process $\mathcal{X}$; this occurs with probability $\lambda(x_j)\,dt_j\,dx_j$; and (ii) during the time interval (t_{j-1},t_j), no events of magnitude greater or equal to the value x_j take place; this occurs with probability $\exp\{-(t_j - t_{j-1})\Lambda(x_j)\}$. Multiplying these probabilities together yields (33).

We note that (33) can be written, alternatively, in the form

$$\begin{aligned}\mathbf{P}(\mathcal{R}(dt_j \times dx_j) &= 1\,; j = 1,\ldots,n)\\ &= \prod_{j=1}^n \exp\{-t_j(\Lambda(x_j) - \Lambda(x_{j+1}))\}\lambda(x_j)\,dt_j\,dx_j,\end{aligned} \tag{34}$$

where x_{n+1} is set to equal $+\infty$.

Now, integrating (33) over $x_1 < x_2 < \cdots < x_n$ yields

$$\mathbf{P}(\mathcal{R}_{\text{time}}(dt_j) = 1\,; j = 1,\ldots,n) = \prod_{j=1}^n \frac{dt_j}{t_j}, \tag{35}$$

and integrating (34) over $t_1 < t_2 < \cdots < t_n$ yields

$$\mathbf{P}(\mathcal{R}_{\text{value}}(dx_j) = 1\,; j = 1,\ldots,n) = \prod_{j=1}^n \frac{\lambda(x_j)\,dx_j}{\Lambda(x_j)}. \tag{36}$$

Hence, using (5), we can conclude the following:

PROPOSITION 3: *Let $\mathcal{X}$ be an event process with characteristic $\Lambda(x)$. Then the following hold:*

https://doi.org/10.1017/S0269964805050163 Published online by Cambridge University Press

(i) *The set of record times* $\mathcal{R}_{\text{time}}$ *of* $\mathcal{X}$ *is a temporal Poisson point process with rate* $1/t$.
(ii) *The set of record values* $\mathcal{R}_{\text{value}}$ of $\mathcal{X}$ *a is Poisson point process (on* $\mathbb{R}$*) with rate* $\lambda(x)/\Lambda(x)$.

In particular, for the Fréchet, Weibull, and Gumbel cases the rate $\lambda(x)/\Lambda(x)$ of the Poisson point process of record values $\mathcal{R}_{\text{value}}$ is as follows:

Fréchet:	α/x,	$x > 0$;
Weibull:	$\alpha/\lvert x\rvert$,	$x < 0$;
Gumbel:	1,	x real.

Several remarks are necessary:

(i) We emphasize that the record set itself $\mathcal{R}$ is *not* a Poisson point process on $[0,\infty) \times \mathbb{R}$ since the probability $\mathbf{P}(\mathcal{R}(dt_j \times dx_j) = 1\,; j = 1,\ldots,n)$ given in (34) does *not* admit the multiplicative form (5).
(ii) Equation (35) is the continuous-time counterpart of Rényi's record theorem [18]. Rényi's theorem asserts that if $\{\xi_j\}_{j=1}^n$ is an i.i.d. sequence of random variables and E_j is the event $\{\xi_j$ is a record value$\}$, then the events $\{E_j\}_{j=1}^n$ are independent and $\mathbf{P}(E_j) = 1/j$.
(iii) Results analogous to those given in (35) and (36) for i.i.d. random sequences can be found in [3, Chap. 5].
(iv) Recall the waiting time $L(t)$ defined in Subsection 3.2: $L(t)$ is the length of the time period elapsing from time t until the *first* occurrence of a record event *after* time t. Note that $\{L(t) > l\}$ if and only if the set of record times $\mathcal{R}_{\text{time}}$ has no points on the interval $[t, t+l]$. Hence, Proposition 3 yields that

$$\mathbf{P}(L(t) > l) = \exp\left\{-\int_t^{t+l} \frac{1}{u}\, du\right\} = \frac{1}{1 + l/t},$$

reaffirming (16).

7. CONCLUSIONS

Motivated by the fact that the classic EVT undertakes a discrete-time approach, whereas most "real-world" systems are usually continuous time, our aim in this work was *to introduce a simple and parsimonious model, counterpart to the discrete-time i.i.d. model, for the study of extremes in continuous time*.

To that end, we considered a generic continuous-time system in which events of random magnitudes arrive stochastically—following time-homogeneous Poisson point process dynamics on $[0,\infty) \times \mathbb{R}$. The (monotone decreasing) function $\Lambda(x)$ was used to denote the Poissonian rate at which events of magnitude greater than x occur and was assumed to satisfy $\lim_{x\to-\infty} \Lambda(x) = +\infty$ and $\lim_{x\to+\infty} \Lambda(x) = 0$. This ensured the following: (i) the occurrences of the events are everywhere dense

https://doi.org/10.1017/S0269964805050163 Published online by Cambridge University Press

on the time axis (i.e., there are countably many events occurring on any time interval), whereas (ii) events greater than any given level occur discretely (i.e., there are only finitely many such events on any time interval).

This Poissonian model turned out to naturally accommodate the study and investigation of extremes in continuous time:

- *the maximum* $X_1(t)$: the magnitude of the greatest event occurring during the time interval $[0, t]$;
- *the* nth *"runner-up"* $X_{n+1}(t)$: the magnitude of the $(n+1)$st greatest event occurring during the time interval $[0, t]$;
- *the "top n"* $(X_1(t), \ldots, X_n(t))$: the vector of magnitudes of the n greatest events occurring during the time interval $[0, t]$;
- *the set of record times* R_{time}: the time epochs at which the record events occurred;
- *the set of record values* R_{value}: the values of the record events.

Furthermore, for any fixed time t, we explored the sequence of order statistics $\{X_n(t)\}_{n=1}^{\infty}$—viewed as a stochastic process indexed by the parameter n:

- *Structure*: We discovered that the increasing sequence of points $\{\Lambda(X_n(t))\}_{n=1}^{\infty}$ forms a standard Poisson Process with rate t.
- *Simulation*: We devised a markedly simple algorithm for the simulation of the sequence $\{X_n(t)\}_{n=1}^{\infty}$.
- *Hierarchy*: We discovered that the ratios $\{\Lambda(X_n(t))^n/\Lambda(X_{n+1}(t))^n\}_{n=1}^{\infty}$ are independent and uniformly distributed on the unit interval.

Throughout, emphasis was focused on three special cases: (i) Fréchet $\Lambda(x) = x^{-\alpha}$ $(x > 0)$; (ii) Weibull $\Lambda(x) = |x|^{\alpha}$ $(x < 0)$; (iii) Gumbel $\Lambda(x) = \exp\{-x\}$ (x real). For these cases, the resulting maximum distributions are governed respectively by the Fréchet, Weibull, and Gumbel probability laws. These laws are of major importance since they are the universal "stable laws" of the classic EVT—the only possible maxima scaling limits.

We hope that researchers in the engineering and informational sciences, when having to deal with extreme events in continuous-time systems, would find use in the modeling approach and results presented in this work.

Acknowledgments

The author would like to thank the anonymous referee for his helpful and constructive comments.

References

1. Bertoin, J. (1996). *Lévy processes*. Cambridge: Cambridge University Press. Bertoin, J. (1999). *Subordinators: Examples and applications*. Lecture Notes in Mathematics 1717. New York: Springer-Verlag.
2. Eliazar, I. (2005). On geometric record times. *Physica A* 348: 181–198.
3. Embrechts, P., Kluppelberg, C., & Mikosch, T. (1997). *Modelling extremal events for insurance and finance*. Springer-Verlag.

4. Embrechts, P. & Omey, E. (1983). On subordinated distributions and record processes. *Mathematical Proceedings of the Cambridge Philosophical Society* 93: 339–353.
5. Evans, M., Hastings, N., & Peacock, B. (1993). *Statistical distributions*, 2nd ed. New York: Wiley.
6. Feller, W. (1971). *An introduction to probability theory and its applications*, Vol. 2, 2nd ed. New York: Wiley.
7. Fisher, R.A. & Tippett, L.H.C. (1928). Limiting forms of the frequency distribution of the largest and smallest member of a sample. *Proceedings of the Cambridge Philosophical Society* 24: 180–190.
8. Fréchet, M. (1927). Sur la loi de probabilité de l'écart maximum. *Ann. Soc. Polon. Math. Cracovie* 6: 93–116.
9. Galambos, J. (1987). *Asymptotic theory of extreme order statistics*, 2nd ed. New York: Krieger.
10. Gaver, D.P. (1976). Random record models. *Journal of Applied Probability* 13: 538–547.
11. Gnedenko, B. (1943). Sur la distribution limite du terme maximum d'une serie aleatoire. *Annals of Mathematics* 44: 423. Translated and reprinted in S. Kotz & N.L. Johnson, eds. (1992). *Breakthroughs in statistics I*. Berlin: Springer-Verlag, pp. 195–225.
12. Gumbel, E.J. (1958). *Statistics of extremes*. New York: Columbia University Press.
13. Ibragimov, I.A. & Linnik, Yu.V. (1971). *Independent and stationary sequences of random variables*. Groningen: Walters-Noordhoff.
14. Kingman, J.F.C. (1993). *Poisson processes*. Oxford: Oxford University Press.
15. Kotz, S. & Nadarajah, S. (2000). *Extreme value distributions*. London: Imperial College Press.
16. Levikson, B., Rolski, T., & Weiss, G. (1999). On the Poisson hyperbolic staircase. *Probability in the Engineering and Informational Sciences* 13: 11–32.
17. Reiss, R.D. (1993). *A course on point processes*. New York: Springer-Verlag.
18. Rényi, A. (1976). On outstanding values of a sequence of observations. In: P. Turan (ed.), *Selected papers of A. Rényi*, Vol. 3. Budapest: Akadémiai Kiadó.
19. Thomas, M. & Reiss, R.D. (2001). *Statistical analysis of extreme values*. Boston: Birkhauser.
20. von Bortkiewicz, L. (1922). Variationsberiteund mittlerer Fehler, *Sitzungsber. Berlin Math. Ges.* 21: 3.
21. von Mises, R. (1936). La distribution de la plus grande de *n* valuers. *Rev. Math. Union Interbalcanique* 1: 141–160. Translated and reprinted in (1964). *Selected papers of Richard von Mises II*. Providence, RI: American Mathematical Society, pp. 271–294.
22. Westcott, M. (1977). The random record model. *Proceedings of the Royal Society, London, A* 356: 529.

APPENDIX

We introduce an order-preserving transformation of event processes and use it to prove Proposition 2.

A.1. An Order-Preserving Transformation

Given a point process $\mathcal{X}$ on $[0,\infty) \times I$ and a monotone-increasing bijection $\phi : I \to J$ (I and J being subintervals of the real line $\mathbb{R}$), consider the transformation

$$\Phi : \mathcal{X} \mapsto \mathcal{Y} = \{(t, \phi(x))\}_{(t,x)\in\mathcal{X}}; \tag{A.1}$$

that is, the point (t,x) of $\mathcal{X}$ is mapped by the transformation Φ to the point $(t,\phi(x))$ of $\mathcal{Y}$. The map Φ transforms the point process $\mathcal{X}$ (on $[0,\infty) \times I$) to a new point process $\mathcal{Y}$ (on $[0,\infty) \times J$). Furthermore, the map Φ *preserves the order of the order statistics*: If $\{X_n(t)\}_{n=1}^{\infty}$ is the sequence of order statistics of the process $\mathcal{X}$ and $\{Y_n(t)\}_{n=1}^{\infty}$ is the sequence of order statistics of the process $\mathcal{Y}$, then

https://doi.org/10.1017/S0269964805050163 Published online by Cambridge University Press

$$Y_n(t) = \phi(X_n(t)). \tag{A.2}$$

Now, if $\mathcal{X}$ is a Poisson point process with rate $r_{\mathcal{X}}(t,x)$, then standard probabilistic arguments imply that $\mathcal{Y}$ is also a Poisson point process and its rate is $r_{\mathcal{Y}}(t,y) = r_{\mathcal{X}}(t,\phi^{-1}(y))/\phi'(\phi^{-1}(y))$. In particular, if $\mathcal{X}$ is a sample process with characteristic $\Lambda_{\mathcal{X}}(x)$, then $\mathcal{Y}$ is a new sample process with characteristic

$$\Lambda_{\mathcal{Y}}(y) = \Lambda_{\mathcal{X}}(\phi^{-1}(y)). \tag{A.3}$$

Equation (A.3) can also be derived directly by combining together (A.2) and (14).

Hence, given a pair of sample processes—$\mathcal{X}$ with characteristic $\Lambda_{\mathcal{X}}(x)$ and $\mathcal{Y}$ with characteristic $\Lambda_{\mathcal{Y}}(y)$—the map Φ induced by the monotone-increasing bijection

$$\phi(x) = \Lambda_{\mathcal{Y}}^{-1}(\Lambda_{\mathcal{X}}(x)) \tag{A.4}$$

transforms the process $\mathcal{X}$ to the process $\mathcal{Y}$.

Equations (A.3) and (A.4) enable us to *transform* between different sample processes while *preserving* the order of their order statistics.

A.2. Proof of Proposition 2

We split the proof into two steps. In the first step, we prove that Proposition 2 holds for the special case of an event process $\mathcal{Y}$ with characteristic $\Lambda_{\mathcal{Y}}(y) = 1/y$, $y > 0$. In the second step, we use the order-preserving transformation of Subsection A.1 to validate Proposition 2 for any event process $\mathcal{X}$.

We remark that the running maximum process $(Y_1(t))_{t\geq 0}$ (associated with the event process $\mathcal{Y}$) is analogous, albeit not identical, to the "Poisson hyperbolic staircase process" introduced and studied in [16].

Step 1: Fix an integer n. The characteristic of the event process $\mathcal{Y}$ is $\Lambda_{\mathcal{Y}}(y) = 1/y$ $(y > 0)$ and, hence, (21) implies that the joint probability density function of the vector of order statistics $(Y_1(t),\ldots,Y_{n+1}(t))$ is given by

$$f_{n+1}(y_1,\ldots,y_{n+1}) = \frac{t^{n+1}\exp\{-t/y_{n+1}\}}{y_1^2 y_2^2 \ldots y_n^2 y_{n+1}^2}.$$

Now,

$$\begin{aligned}
&\mathbf{P}\left(\frac{Y_2(t)}{Y_1(t)} \leq u_1, \frac{Y_3(t)}{Y_2(t)} \leq u_2, \ldots, \frac{Y_{n+1}(t)}{Y_n(t)} \leq u_n\right) \\
&= \underset{\frac{y_2}{y_1}\leq u_1 \quad \frac{y_3}{y_2}\leq u_2 \;\ldots\; \frac{y_{n+1}}{y_1}\leq u_n}{\iint\ldots\int} f_{n+1}(y_1,y_2,\ldots,y_{n+1})\,dy_1\,dy_2\ldots dy_{n+1} \\
&= \int_{y_{n+1}=0}^{\infty}\int_{y_n=y_{n+1}/u_n}^{\infty}\cdots\int_{y_2=y_3/u_2}^{\infty}\int_{y_1=y_2/u_1}^{\infty} f_{n+1}(y_1,y_2,\ldots,y_{n+1})\,dy_1\,dy_2\ldots dy_{n+1} \\
&\quad\vdots \\
&= u_1^1 u_2^2 \ldots u_n^n \int_{y_{n+1}=0}^{\infty} \frac{t^{n+1}}{n!}\,\frac{\exp\{-t/y_{n+1}\}}{y_{n+1}^{n+2}}\,dy_{n+1} \\
&= u_1^1 u_2^2 \ldots u_n^n.
\end{aligned}$$

https://doi.org/10.1017/S0269964805050163 Published online by Cambridge University Press

This, in turn, implies that ratios $\{Y_{j+1}(t)/Y_j(t)\}_{j=1}^n$ are *independent* random variables governed by the *Beta distribution*:

$$\mathbf{P}\left(\frac{Y_{j+1}(t)}{Y_j(t)} \le u\right) = u^j.$$

Step 2: Let $\mathcal{X}$ be an arbitrary event process with characteristic $\Lambda_{\mathcal{X}}(x)$ and let $\mathcal{Y}$ be an event process with characteristic $\Lambda_{\mathcal{Y}}(y) = 1/y$, $y > 0$. Equation (A.4) implies that the order-preserving map Φ transforming the process $\mathcal{X}$ to the process $\mathcal{Y}$ is induced by the monotone-increasing bijection

$$\phi(x) = \frac{1}{\Lambda_{\mathcal{X}}(x)}$$

(or, equivalently, using (A.3): if $\phi(x) = 1/\Lambda_{\mathcal{X}}(x)$, then $\Lambda_{\mathcal{Y}}(y) = 1/y$). Exploiting the fact that Φ is order-preserving, (A.2) implies that

$$\frac{Y_{j+1}(t)}{Y_j(t)} = \frac{\phi(X_{j+1}(t))}{\phi(X_j(t))} = \frac{\Lambda_{\mathcal{X}}(X_j(t))}{\Lambda_{\mathcal{X}}(X_{j+1}(t))}.$$

Hence, using Step 1, we obtain that the ratios $\{\Lambda_{\mathcal{X}}(X_j(t))/\Lambda_{\mathcal{X}}(X_{j+1}(t))\}_{j=1}^n$ are *independent* random variables governed by the *Beta distribution*:

$$\mathbf{P}\left(\frac{\Lambda_{\mathcal{X}}(X_j(t))}{\Lambda_{\mathcal{X}}(X_{j+1}(t))} \le u\right) = u^j.$$

Since the choice of n was arbitrary, the proof of Proposition 2 is complete.