Extremes of events with heavy-tailed inter-arrival times

Peter Straka*, Smarak Nayak
UNSW Sydney, Australia - p.straka@unsw.edu.au

Abstract

Heavy-tailed inter-arrival times are a signature of "bursty" dynamics, and have been observed in financial time series, earthquakes, solar flares, neuron spike trains and many social, technological and economic phenomena. The study of extremes in bursty time series is important for risk assessment, resource management and service allocation. "Standard" Extreme Value Statistics assumes evenly spaced or Poisson process arrivals, and hence cannot model bursty dynamics. We propose a renewal process with power-law, infinite mean inter-arrival times as a model for bursty dynamics, and assume i.i.d. event magnitudes at the renewal times. This results in a so-called "Max-Renewal process" or "Continuous Time Random Maxima process". The renewal process assumption is rarely satisfied on the scale of individual observations, but often so for the times between events larger than a given (high) threshold. These threshold exceedance times are, due to geometric sum-stability, attracted to a Mittag-Leffler distribution. We show that as the threshold height is increased, the Mittag-Leffler shape parameter stays constant, while the scale parameter grows like a power-law. This theoretical result is confirmed by observing the same behaviour on datasets for earthquakes, solar flares and financial losses. We discuss approaches to fit model parameters and uncertainty due to threshold selection.

Keywords: renewal process; geometric stability; extreme value statistics; bursts.

1. Introduction

Time series displaying inhomogeneous behaviour have received strong interest in the recent statistical physics literature, [Bar05], and have been observed in the context of earthquakes, sunspots, neuronal activity and human communication, see [KKBK12, VTK13] for a list of references. Such time series exhibit high activity in some 'bursty' intervals, which alternate with other, quiet intervals. Although several mechanisms are plausible explanations for bursty behaviour (most prominently self-exciting point processes [Haw71]), there seems to be one salient feature which very typically indicates the departure from temporal homogeneity: a heavy-tailed distribution of waiting times. A simple renewal process with heavy-tailed waiting times captures these dynamics. For many systems, the renewal property is appropriate, as the dynamics do not change visibly if the waiting times are randomly reshuffled [KKBK12].

When a magnitude can be assigned to each event in the renewal process, such as for earthquakes, sun flares, neuron voltages or the impact of an email, it is natural to consider the consequences of *extreme* events. Thus two questions that arise are: When will the next extreme event occur? And, how large will it be? A probabilistic extreme value model which assumes the renewal property can be used to answer these questions. One such model has been studied under the names "Continuous Time Random Maxima process" (CTRM), see e.g. [MS08], "Max-Renewal process", see e.g. [ST04], and "Shock process", e.g. [SS83]. In this work, we discuss ideas for methods of inference for this type of model, a problem which has seemingly received little attention by the statistical community.

2. CTRMs

The Continuous Time Random Walk (CTRW) has been a highly successful model for anomalous diffusion in the past two decades [MK00, HTSL10], likely due to its tractable and flexible scaling properties. The stochastic process we study here is conceptually very close to the CTRW, since essentially the jumps J_k are reinterpreted as magnitudes, and instead of the cumulative sum, one tracks the cumulative maximum.

Similarly tractable scaling properties apply to the CTRM, and many of the results we present for the CTRM have been derived as extensions to similar results for the CTRW [BŠ14, MS08, HS17].

Definition 1. *Assume i.i.d. inter-arrival times* $W_k > 0$ *and event magnitudes* $J_k \in [x_L, x_R]$*, where* $x_L, x_R \in [-\infty, \infty]$*. Define the renewal process associated with the* W_k *as*

$$N(t) = \max\{n \in \mathbb{N} : S(n) \leq t\}, \tag{2}$$

and let $M(n) := \bigvee_{k=1}^{n} J_k$ *denote the running maximum. Then the process*

$$V(t) = M(N(t)) = \bigvee_{k=1}^{N(t)} J_k, \quad t \geq 0. \tag{3}$$

is called a CTRM (Continuous Time Random Maxima) process. For convenience, we set $M(0) = x_L$*, and* $\max\{\} = 0$.

It is clear from the definition that the sample paths of $N(t)$ and $V(t)$ are right-continuous with left-hand limits. If W_k is interpreted as the time *leading up* to the event with magnitude J_k, then $V(t)$ is the largest magnitude observed until time t. The alternative case where W_k represents the inter-arrival time *following* J_k is termed "second type" (in the shock model literature, [SS83]) or OCTRM (overshooting CTRM) [HS17], and the largest magnitude up to time t is then given by

$$\tilde{V}(t) = \bigvee_{k=1}^{N(t)+1} J_k, \quad t \geq 0. \tag{4}$$

Finally, the CTRM model is called *coupled* when W_k and J_k are not independent. We focus on the uncoupled case, for which it can be shown that the processes $V(t)$ and $\tilde{V}(t)$ have the same limiting distributions at large times [HS16]. When drawing inference on $V(t)$ it is natural to be concerned with both the timing and magnitude of extreme events. By definition extreme events are those of large magnitude and therefore must exceed a high threshold ℓ. Thus we seek the distribution of the following quantities:

Definition 5. *The* ***exceedance time*** *resp.* ***exceedance*** *of level* $\ell \in [x_L, x_R]$ *are the random variables*

$$T_\ell = \inf\{t : V(t) > \ell\}, \qquad X_\ell = V(T_\ell) - \ell.$$

With a few towards statistics, T_ℓ and X_ℓ are both directly observable in a dataset: Figure 1 shows magnitudes J_k at their arrival times $W_1 + \ldots + W_k$. Given a threshold (dashed line), this defines a sequence of exceedances $\{X_{\ell,k}\}_{k\in\mathbb{N}}$ (lengths of red lines) and exceedance times $\{T_{\ell,k}\}_{k\in\mathbb{N}}$ (distances between adjacent blue circles), which will be used for inference later.

Lemma 6. *Let* $V(t)$ *be an uncoupled CTRM process, and let* $\ell \in [x_L, x_R]$.

1. *The exceedances* $\{X_{\ell,k}\}_{k\in\mathbb{N}}$ *and exceedance times* $\{T_{\ell,k}\}_{k\in\mathbb{N}}$ *are both i.i.d.*
2. *For each* $k \in \mathbb{N}$*,* $X_{\ell,k}$ *and* $T_{\ell,k}$ *are independent.*

Proof. This is a direct consequence of the i.i.d. property of W_k and J_k. □

A standard result from extreme value theory provides the (asymptotic) distribution of X_ℓ; see Theorem 7 below, which we include for completeness. We can hence devote the remainder of this text to the estimation of the distribution of T_ℓ.

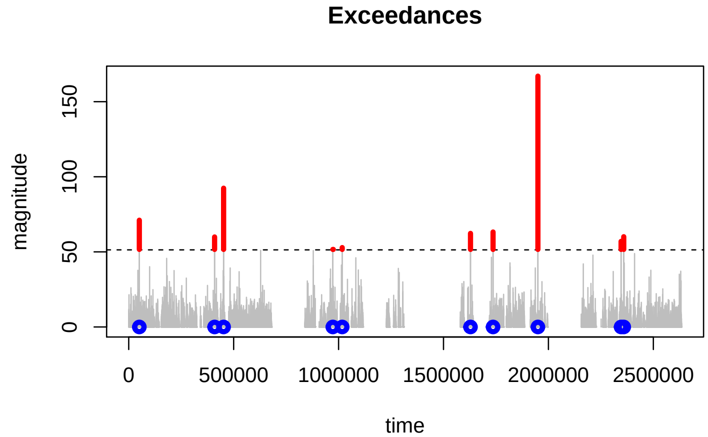

Figure 1: A thresholded CTRM process.

Theorem 7 (e.g. [BGST06])**.** *Suppose the distribution of J_k is continuous. Then there exist non-decreasing norming functions $a(n)$ and $d(n)$ such that as $n \to \infty$, the running maximum $M(n)$ converges weakly to a **Generalized Extreme Value Distribution** with shape parameter $\xi \in \mathbb{R}$:*

$$[M(n) - d(n)]/a(n) \stackrel{d}{\to} A, \qquad \mathbb{P}(A \le z) = G(z) = \exp\left(-[1+\xi z]^{-1/\xi}\right),$$

defined for z such that $1 + \xi z > 0$, and defined for $\xi = 0$ via continuity. We write $\mathrm{GEV}(\xi, \mu, \sigma)$ for the probability distribution of the random variable $\sigma A + \mu$. Moreover, asymptotically for large $\ell \uparrow x_R$,

$$\mathbb{P}(X_\ell > y) = \mathbb{P}(J_1 - \ell > y | J_1 > \ell) \approx (1 + \xi y/\tilde{\sigma})^{-1/\xi} =: 1 - H(y)$$

*where $\tilde{\sigma} := \sigma + \xi(\ell - \mu)$, $y > 0$ and $1 + \xi y/\tilde{\sigma} > 0$. A distribution with CDF $H(y)$ is said to be from the **Generalised Pareto family** $GP(\xi, \bar{\sigma})$.*

3. Scaling limits

Waiting times. In this text, we assume that the waiting times W_k have a tail parameter $\beta \in (0, 1)$, i.e. $\mathbb{P}(W_1 > t) \sim L(t)t^{-\beta}$ as $t \uparrow \infty$ for some slowly varying function $L(t)$. (We write $f(t) \sim g(t)$ if their quotient converges to 1.) The law of the W_k then lies in the domain of attraction of a stable law, in the sense that the limit

$$(W_1 + \ldots + W_n)/b(n) \stackrel{d}{\longrightarrow} D, \quad n \to \infty \tag{8}$$

exists, for a regularly varying scaling function $b(n) = n^{1/\beta}L(n) \in \mathrm{RV}_\infty(1/\beta)$, where $L(n)$ is slowly varying. The limit is then a positively skewed stable distribution, whose scale parameter is 1 is $b(n)$ is chosen accordingly; that is, $\mathbb{E}[\exp(-sD)] = \exp(-s^\beta)$. Moreover, the following functional limit theorem holds:

$$(W_1 + \ldots + W_{\lfloor ct \rfloor})/b(c) \stackrel{d}{\longrightarrow} D(t), \quad c \to \infty \tag{9}$$

with convergence in the Skorokhod J_1 topology. The limit $D(t)$ is a stable subordinator, i.e. an increasing Lévy process with Laplace transform $\exp(-ts^\beta)$.
It is well known (see e.g. [MS04]) that the renewal process then satisfies the functional limit

$$N(ct)/\tilde{b}(c) \stackrel{d}{\longrightarrow} E(t) = \inf\{r : D(r) > t\}, \quad c \to \infty \tag{10}$$

for a scaling function $\tilde{b}(c)$ which is asymptotically inverse to $b(c)$, in the sense of [Sen76, p.20]:

$$b(\tilde{b}(c)) \sim c \sim \tilde{b}(b(c)). \tag{11}$$

Note that $\tilde{b} \in \mathrm{RV}_\infty(\beta)$ [MS04]. The limit process $E(t)$ is called the *inverse* stable subordinator [MS13].

Event magnitudes. The extremal limit theorem allows for an extension to a functional limit: assume that the norming sequences $a(n)$ and $d(n)$ are as in Theorem 7. Then

$$[M(\lfloor ct \rfloor) - d(c)]/a(c) \xrightarrow{d} A(t), \quad c \to \infty.$$

Convergence is in Skorokhod's J_1 topology, and the limit process $A(t)$ is an *extremal process*, with finite-dimensional distributions given by

$$\mathbb{P}(A(t_i) \le x_i, 1 \le i \le d) = F_A(\wedge_{i=1}^d x_i)^{t_1} F_A(\wedge_{i=2}^d x_i)^{t_2 - t_1} \dots F_A(x_d)^{t_d - t_{d-1}},$$

with $F_A(x)$ being a GEV distribution.

CTRM limit. A CTRM is running-maximum process $M(n)$, time-changed by the renewal process $n = N(t)$. Accordingly, its scaling limit results from the time-change of its corresponding limit processes:

Theorem 12. *[MS08] The CTRM process $V(t) = M(N(t))$ satisfies the following functional scaling limit in the Skorokhod J_1 topology:*

$$[V(ct) - d(\tilde{b}(c))]/a(\tilde{b}(c)) \xrightarrow{d} A(E(t)), \quad c \to \infty.$$

The distribution of the *hitting time* of a level ℓ by the limit process $A(E(t))$ has been found as well:

Theorem 13. *Let F_A be the CDF of a GEV distribution, and let $A(t)$ be the corresponding extremal process. For a given threshold ℓ^* in the support of F_A, the hitting time $B(\ell^*) = \inf\{t : A(E(t)) > \ell^*\}$ follows a Mittag-Leffler distribution:*

$$B(\ell^*) \sim \mathrm{ML}\left(\beta, (-\log F_A(\ell^*))^{-1/\beta}\right)$$

Recall that for $\beta \in (0,1)$, a standard Mittag-Leffler random variable Y is positive with Laplace transform $\mathbb{E}[\exp(-sY)] = 1/(1+s^\beta)$. For $\sigma > 0$, we then write $\mathrm{ML}(\beta, \sigma)$ for the distribution of σY.

4. Inference on Exceedance Times

Distribution of exceedance times. We now proceed to the main result of this paper, which is to derive a statistical inference procedure for the exceedance times T_ℓ based on Theorem 13. Using $c \approx b(\tilde{b}(c))$, writing $n = \tilde{b}(c)$, and substituting $t \to t/b(n)$, Theorem 12 implies

$$V(t) \approx a(n) A(E(t/b(n))) + d(n).$$

Hence

$$T_\ell > t \Leftrightarrow V(t) \le \ell \Leftrightarrow A(E(t/b(n))) \le \frac{\ell - d(n)}{a(n)} \Leftrightarrow \xi_{\ell^*} > t/b(n)$$

where $\ell^* = (\ell - d(n))/a(n)$. This shows that, approximately, $T_\ell \stackrel{d}{=} b(n)\xi_{\ell^*}$, or, using Theorem 13,

$$T_\ell \sim \mathrm{ML}\left(\beta, b(n)(-\log F_A(\ell^*))^{-1/\beta}\right)$$

By extremal limit theory,

$$-\log F_A(\ell^*) \approx -\log F_J(\ell)^n = -n \log(1-p) \approx np$$

where $p = 1 - F_J(\ell) = 1 - \mathbb{P}(J_k \le \ell)$, and hence we may write

$$T_\ell \sim \mathrm{ML}(\beta, p^{-1/\beta} L(n)), \tag{14}$$

recalling from above that $b(n) = n^{1/\beta} L(n)$.

Estimation algorithm. A natural estimator for p is $|\{k : J_k > \ell\}|/n$, i.e. the fraction of magnitudes which exceed ℓ. We note that the tail parameter β does not depend on the threshold level ℓ. Inspired by the POT (points over threshold) method and using Equation (14), we hence suggest the following estimation procedure for the distribution of T_ℓ:

1. For varying thresholds $\ell \uparrow x_R$, extract datasets of exceedance times $\{T_{\ell,k}\}_k$.
2. For each ℓ, fit a Mittag-Leffler distribution to $\{T_{\ell,k}\}_k$, yielding the estimates $\{\hat\beta(\ell)\}_\ell$ and $\{\hat\sigma(\ell)\}_\ell$.
3. Plot ℓ vs. $\hat\beta(\ell)$, and look for regions where the plot looks stable. Choose $\hat\beta$ from the high-ℓ end of this region.
4. Approximating $p \approx |\{k : J_k > \ell\}|/n$, plot ℓ vs. $p^{1/\hat\beta}\hat\sigma(\ell)$, and look for regions where the plot looks stable. Choose $\hat\sigma_0$ from the high-ℓ end of this region.

One can then estimate the distribution of the exceedance times for the threshold ℓ via a $\mathrm{ML}(\hat\beta, p^{-1/\hat\beta}\hat\sigma_0)$. Figure 2 shows the diagnostic plots from Steps 3 & 4, for a dataset of magnitudes and occurrence times of earthquakes in the Coral Sea, obtained from `http://www.seismicportal.eu`. Estimates of the Mittag-Leffler parameters were calculated using the log-moments method [Cah13]. We suggest the estimates $\hat\beta \approx 0.5$, $\hat\sigma \approx 1.5 \times 10^7$ seconds, approximately 0.48 years.

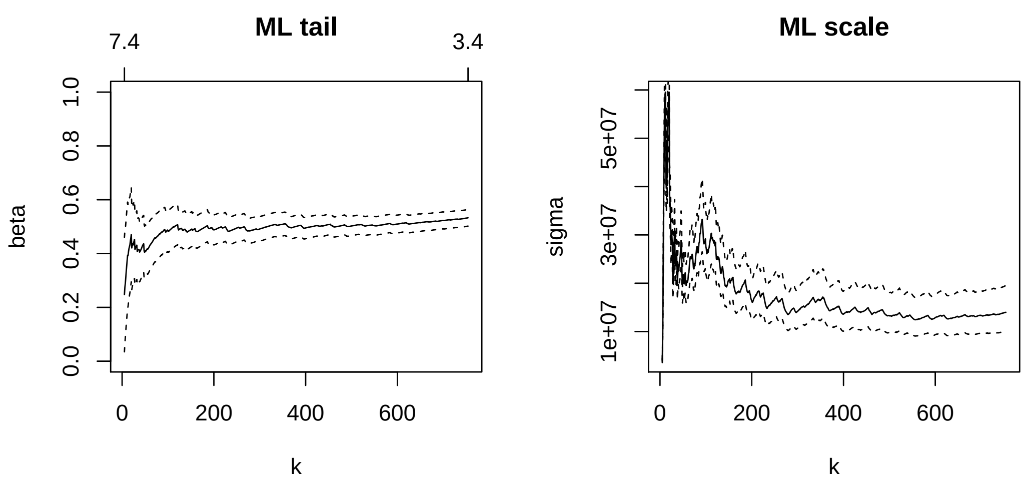

Figure 2: Stability Plots for an earthquake dataset. The x-axis k denotes the number of obervations above the threshold ℓ.

5. Conclusion

Standard extreme value theory [BGST06] provides classical methods for the statistical fitting of extreme value distributions. If the times between events are light-tailed, this theory entails that the times between large events are exponentially distributed. We have extended this theory, in particular the POT (points over threshold) method, to apply to heavy-tailed inter-event times, generalizing the exponential distribution to a Mittag-Leffler distribution (note that the exponential distribution is a special case of Mittag-Leffler, for $\beta = 1$).

For brevity, we have omitted plots of the autocorrelation function for the sequence of waiting times, which shows that the renewal assumption is not grossly disobeyed, and Q-Q plots for the quality of the fit of the Mittag-Leffler distribution.

Finally, there are multiple avenues for a refinement of this method: How can the uncertainty of the parameters be quantified, and how does it propagate to the uncertainty of predictions? How can uncertainty in the choice of threshold be quantified, again with a view towards predictions and the estimation of risk? And finally, we deem it possible to extend the above theory to semi-heavy-tails, where waiting times have infinite variance but finite mean.

Acknowledgements. P. Straka was supported by the Australian Research Councils Discovery Early Career Research Award DE160101147.

References

[Bar05] Albert László Barabási. The origin of bursts and heavy tails in human dynamics. *Nature*, 435(May):207–211, 2005.

[BGST06] Jan Beirlant, Yuri Goegebeur, Johan Segers, and Jozef Teugels. *Statistics of extremes: theory and applications.* John Wiley & Sons, 2006.

[BŠ14] Bojan Basrak and Drago Špoljarić. Extremal behaviour of random variables observed in renewal times. jun 2014.

[Cah13] Dexter O. Cahoy. Estimation of Mittag-Leffler Parameters. *Commun. Stat. - Simul. Comput.*, 42(2):303–315, feb 2013.

[Haw71] Alan G Hawkes. Point spectra of some mutually exciting point processes. *J. R. Stat. Soc. Ser. B*, pages 438–443, 1971.

[HS16] Katharina Hees and Hans-Peter Scheffler. On joint sum/max stability and sum/max domains of attraction. pages 1–31, jun 2016.

[HS17] Katharina Hees and Hans-Peter Scheffler. Coupled Continuous Time Random Maxima. pages 1–24, feb 2017.

[HTSL10] B.I. Henry, T.A.M. Langlands, and Peter Straka. An introduction to fractional diffusion. *Complex Phys. Biophys. Econophysical Syst.* World Scientific, 2010.

[KKBK12] Márton Karsai, Kimmo Kaski, Albert László Barabási, and János Kertész. Universal features of correlated bursty behaviour. *Sci. Rep.*, 2, 2012.

[MK00] Ralf Metzler and Joseph Klafter. The random walk's guide to anomalous diffusion: a fractional dynamics approach. *Phys. Rep.*, 339(1):1–77, dec 2000.

[MS04] Mark M Meerschaert and Hans-Peter Scheffler. Limit Theorems for Continuous-Time Random Walks with Infinite Mean Waiting Times. *J. Appl. Probab.*, 41(3):623–638, sep 2004.

[MS08] Mark M Meerschaert and Stilian A Stoev. Extremal limit theorems for observations separated by random power law waiting times. *J. Stat. Plan. Inference*, 139(7):2175–2188, jul 2008.

[MS13] Mark M Meerschaert and Peter Straka. Inverse Stable Subordinators. *Math. Model. Nat. Phenom.*, 8(2):1–16, apr 2013.

[Sen76] E Seneta. *Regularly Varying Functions*, volume 508 of *Lecture Notes in Mathematics.* Springer-Verlag, Berlin, 1976.

[SS83] J. George Shanthikumar and Ushio Sumita. General shock models associated with correlated renewal sequences. *J. Appl. Probab.*, 20(3):600–614, 1983.

[ST04] Dmitrii S Silvestrov and Jozef L. Teugels. Limit theorems for mixed max-sum processes with renewal stopping. *Ann. Appl. Probab.*, 14(4):1838–1868, nov 2004.

[VTK13] Szabolcs Vajna, Bálint Tóth, and János Kertész. Modelling bursty time series. *New J. Phys.*, 15(10):103023, oct 2013.