Extreme Events: Dealing With Dependence

Johan Segers[†]

October 16, 2002

Abstract. Classical extreme value theory for stationary sequences of random variables can up to a large extent be paraphrased as the study of exceedances over a high threshold. Much is known about the asymptotic dependence structure between these exceedances, mostly in terms of the extremal index and its various characterizations. Parts of this theory now can be generalized not only to random variables on an arbitrary state space hitting certain failure sets but even to a triangular array of events on an abstract probability space. A coefficient is also introduced to describe the dependence between two such triangular arrays of events. Finite-sample inequalities lead to asymptotic results under rather weak stationarity and mixing conditions. Applications include a sliding-blocks estimator for the probability of no extreme event in a large block of time and an estimator of the suitably generalized extremal index based on the inter-arrival times between extreme events.

AMS Subject Classification: primary 62 G 32; secondary 62 G 05, 62 G 20

Keywords: block maximum; blocks estimator; cluster of exceedances; extremal index; failure set; mixing condition; stationary sequence; threshold exceedance

[†]EURANDOM, P.O. Box 513, NL-5600 MB Eindhoven, The Netherlands (E-mail: *segers@eurandom.tue.nl*). The author is a Post-Doctoral Research Fellow of the Fund for Scientific Research - Flanders (Belgium).

1 Introduction

Many applied sciences require handling events with low probability but large, often disastrous impact. Extreme events form a central issue in financial risk management, premium calculations in reinsurance, the construction of dams and drainage systems, metal fatigue, and many more areas. Of particular interest is the way in which extreme events interact: an unusually stormy day at a particular site may well be followed by another one at the same or a neighboring site; a large drop in a stock index may trigger similar negative movements in the next time period for the same or other financial time series. If extreme-value statistics is already complicated by the fact that about the events it wants to describe there are by definition few observations, even more challenging to model and to make inference on are the possible connections between different extreme events. Which, then, are the principles underlying these dependencies between extremes?

The issue will be treated in a rather abstract setting, whose build-up is conveniently commenced at a concept from classical extreme-value theory. A stationary sequence of random variables $\{X_n : n \geq 1\}$ is said to have extremal index $\theta \in [0,1]$ if for every $\tau > 0$ there exist numbers $\{u_n : n \geq 1\}$ such that $n \Pr(X_n > u_n) \to \tau$ and $\Pr(\max_{i=1,\ldots,n} X_i \leq u_n) \to \exp(-\tau\theta)$ as $n \to \infty$ (Leadbetter 1983). The extremal index θ quantifies the strength of the dependence between the threshold exceedances $\{X_i > u_n\}$, with $\theta = 1$ corresponding to asymptotic independence and $\theta \downarrow 0$ to increasing dependence, showing itself in a tendency for large observations to occur in clusters. Under certain mixing conditions on the $\{X_n\}$, the extremal index arises in at least three other ways: as the reciprocal of the mean size of a cluster of threshold exceedances; as the probability that a threshold exceedance is not followed in the near future by another one; and as the shape parameter in the limit distribution of the inter-exceedance times. All these characterizations motivate different estimators of the extremal index and of properties of clusters of extremes. Together with the tail of the marginal distribution of the X_n, they provide a fairly complete picture of the probabilistic structure of extreme events in the time series.

The concept of extremal index has been generalized to multivariate stationary time series $\{\mathbf{X}_n = (X_n^{(1)}, \ldots, X_n^{(d)}) : n \geq 1\}$. Let order relations in $\mathbb{R}^d$ be taken componentwise and define $\mathbf{M}_n = \max\{\mathbf{X}_n : i = 1, \ldots, n\}$. Consider multivariate thresholds $\mathbf{u}_n = (u_n^{(1)}, \ldots, u_n^{(d)}) \in \mathbb{R}^d$ for which $n \Pr(X_n^{(i)} > u_n^{(i)}) \to \tau^{(i)} \in [0, \infty)$ as $n \to \infty$. If both limits

$$n \Pr(\mathbf{X}_1 \not\leq \mathbf{u}_n) \to \lambda \in (0, \infty) \quad \text{and} \quad \Pr(\mathbf{M}_n \leq \mathbf{u}_n) \to \exp(-\mu) > 0$$

exist, then the multivariate extremal index is defined by $\theta(\boldsymbol{\tau}) = \mu/\lambda$. As the notation suggests, the extremal index depends on $\boldsymbol{\tau} = (\tau^{(1)}, \ldots, \tau^{(d)})$ and thus on the threshold sequence $\{\mathbf{u}_n\}$, although under certain mixing conditions $\theta(\boldsymbol{\tau}) = \theta(c\boldsymbol{\tau})$ for $c > 0$. Theory and practice are much less developed than in the univariate case, see Nandagopalan (1994) and Smith and Weissman (1996).

Some reflection on the definition of the (multivariate) extremal index leads to the observation that the order structure on $\mathbb{R}^d$ is not essential and that it is in fact possible to start from a stationary sequence $\{X_n : n \geq 1\}$ of random variables in an arbitrary measurable space $(S, \mathcal{S})$. The thresholds are replaced by measurable sets $B_n \subset S$ which are such that $n \Pr(X_1 \in B_n) \to \tau > 0$ as $n \to \infty$. The process $\{X_n\}$ is said to have extremal index θ w.r.t. $\{B_n\}$ if

$$\Pr(\forall i = 1, \ldots, n : X_i \notin B_n) \to \exp(-\tau\theta) \quad \text{as } n \to \infty.$$

The special cases $S = \mathbb{R}$ and $B_n = (u_n, \infty)$ or, more generally, $S = \mathbb{R}^d$ and $B_n = (-\boldsymbol{\infty}, \mathbf{u}_n]^c$ lead back to the ordinary (multivariate) extremal index. The sets B_n can be thought of as failure sets, which represent a collection of extreme states for the system represented by the X_n. The extremal index θ describes the strength of dependence between the extreme events $\{X_i \in B_n\}$.

At this stage, it is clear that even the X_i and the B_n can be disposed of, and that, finally, the heart of the matter lies in the extreme events $A_{i,n} = \{X_i \in B_n\}$. In general, then, we will work with a triangular array $\{A_{i,n} : n \geq 1, \, i = 1, \ldots, r_n\}$ of events on an abstract probability space (which may vary with n) and for which every row satisfies a certain stationarity condition. When interest is not in asymptotics but in finite-sample statements, the focus will be on a single row $A_1, \ldots, A_r$.

The set-up and the notations in force are detailed in Section 2. In Section 3 we will investigate how close $\Pr(\bigcap_{i=1}^r A_i^c)$ and $\{\Pr(\bigcap_{i=1}^s A_i^c)\}^{r/s}$ are to each other, in terms of finite-sample inequalities as well as asymptotically. This will be applied in Section 4 to a comparison between two estimators of $\Pr(\bigcap_{i=1}^r A_i^c)$ when the indicators I_i of the A_i are observed. In Section 5 properties of the extremal index of a stationary sequence of real-valued random variables will be shown to hold also for

$$\theta_m = \Pr\left(\bigcap_{i=2}^m A_i^c \,\middle|\, A_1 \right),$$

the conditional probability that an extreme event A_1 is followed by a run of $m-1$ non-extreme events A_i^c. These properties will serve in Section 6 to show consistency of the intervals estimator for the extremal index (Ferro and Segers 2002) in our general set-up. Finally, in Section 7 the framework is extended to a double triangular array $\{(A_{i,n}, B_{i,n}) : n \geq 1, \, i = 1, \ldots, r_n\}$. The conditional probability

$$\theta_{m,n}^{A|B} = \Pr\left(\bigcap_{i=1}^m A_{i,n}^c \,\middle|\, \bigcup_{i=1}^m B_{i,n} \right)$$

of no A-event in a block with a B-event will be shown to be an informative coefficient of dependence between the A-array and the B-array.

If classical extreme-value theory for dependent observations allows generalization to such an abstract setting, then inference procedures are feasible in a wide range of situations involving dependence between rare events. Whatever the set-up, the interactions between extremes eventually obey a simple set of basic laws, readily comprehensible and exploitable.

2 Block-stationary events

We describe the basic concepts and assumptions in the theory to be developed. Fundamental is the following notion of stationarity.

Definition 2.1 *Events $A_1, \ldots, A_r$ on a common probability space are block-stationary if for every $m = 1, \ldots, r-1$ and $j = 1, \ldots, r-m$ we have*

$$\Pr\left(\bigcup_{i=1}^{m} A_{i+j}\right) = \Pr\left(\bigcup_{i=1}^{m} A_i\right).$$

Let $A_1, \ldots, A_r$ be block-stationary events. The fundamental quantity of interest is $p_m = \Pr(\bigcup_{i=1}^m A_{i+j})$ for $m = 1, \ldots, r$ and $j = 0, \ldots, r-m$, that is, the probability that a block of length m witnesses an extreme event. Denote $q_m = 1 - p_m$ and $p = p_1$. Check that for positive integers i and j with $i + j \le r$ we have $p_i \le p_{i+j} \le p_i + p_j$ and $q_{i+j} \le q_i \le q_{i+j} + p_j$.

Mixing conditions will be formulated in terms of

$$\begin{aligned} \alpha_{s,l} &= \max\Bigg\{\left|\Pr\left(\bigcap_{i=u+1}^{v} A_i^c \cap \bigcap_{j=s+v+1}^{s+w} A_j^c\right) - q_{v-u}q_{w-v}\right| : \\ &\qquad u \ge 0, v-u \ge l, w-v \ge l, w+s \le r\Bigg\}, \end{aligned}$$

describing the force of dependence between two blocks of length at least l and separated by a gap of size precisely s (put $\alpha_{s,l} = 0$ if $2l + s > r$). Abbreviate $\alpha_l = \alpha_{l,l}$ and $\bar{\alpha}_l = \max\{\alpha_{s,l} : s = l, \ldots, r\}$.

Assuming without further notice that $p > 0$, we define for $m = 1, \ldots, r$

$$\begin{aligned} \theta_m &= (p_m - p_{m-1})/p \\ &= \Pr\left(\bigcap_{i=2}^{m} A_{i+j}^c \,\middle|\, A_{1+j}\right) = \Pr\left(\bigcap_{i=1}^{m-1} A_{i+j}^c \,\middle|\, A_{m+j}\right), \end{aligned}$$

where $j = 0, \ldots, r-m$ and $p_0 = 0$. In words, θ_m is equal to the probability that an extreme event is not followed by another one in the next $m-1$ time points, and also to the probability that an extreme event is not preceded by another one in the previous $m-1$ time points.

The set-up for asymptotic results will be a triangular array $\{A_{i,n} : n \ge 1,\ i = 1, \ldots, r_n\}$ for which every row $A_{1,n}, \ldots, A_{r_n,n}$ consists of block-stationary events on a common probability space, which may vary with n. The probabilities of interest are $p_{m,n} = \Pr(\bigcup_{i=1}^m A_{i+j,n})$ for $m = 1, \ldots, r_n$ and $j = 0, \ldots, r_n - m$, together with $q_{m,n} = 1 - p_{m,n}$ and $p_n = p_{1,n}$. The corresponding mixing coefficients are $\alpha_{s,l,n}$, $\alpha_{l,n} = \alpha_{l,l,n}$, and $\bar{\alpha}_{l,n} = \max\{\alpha_{s,l,n} : s = l, \ldots, r_n\}$. Assuming that $p_n > 0$, we also set $\theta_{m,n} = (p_{m,n} - p_{m-1,n})/p_n$ for $m = 1, \ldots, r_n$, where $p_{0,n} = 0$. Finally, all asymptotic statements are for $n \to \infty$.

Remark 2.2 The condition that the events $A_1, \ldots, A_r$ are block-stationary is weaker than the assumption that the vector of indicator variables $I_i = I(A_i)$ is strictly stationary. As an example, let $\{Y_n : n \in \mathbb{Z}\}$ be independent random variables with $\Pr(Y_n \le y) = \exp(-1/y)$ for $y > 0$, and let a_i, $i \ge 0$, be non-negative numbers such that $a_i \ge a_{i+1}$ for all $i \ge 0$ and $\sum_{i \ge 0} a_i = 1$. Define the moving-maximum process $\xi_n = \max\{a_i Y_{n-i} : i \ge 0\}$, for $n \ge 1$. The process $\{\xi_n\}$ is stationary with block-maximum distribution $\Pr(\max_{i=1,\ldots,n} \xi_i \le x) = \exp\{-[(n-1)a_0 + 1]/x\}$, for $x > 0$. Now let $\{\xi'_n\}$ be another such moving-maximum process, independent of $\{\xi_n\}$, and with parameters a'_i, $i \ge 0$, where again $a'_i \ge a'_{i+1} \ge 0$ for $i \ge 0$ and $\sum_{i \ge 0} a'_i = 1$. Define $(X_1, X_2, X_3, X_4, \ldots) = (\xi_1, \xi'_1, \xi_2, \xi'_2, \ldots)$. If $a_0 = a'_0$ but $a_i \ne a'_i$ for some $i \ge 1$, then the process $\{X_n\}$ is not stationary. Nevertheless, for all $u > 0$ the events $A_n = \{X_n > u\}$ are block-stationary, although the sequence $I_n = I(A_n)$ is not stationary.

Remark 2.3 The mixing coefficients $\alpha_{s,l}$ were introduced by O'Brien (1987) and lead to mixing conditions that are slightly weaker than Leadbetter's (1974) popular condition D.

3 Big and small blocks

A simple but crucial observation for independent and identically distributed real-valued random variables $\{X_n : n \ge 1\}$ is that the distribution of the sample maximum $M_n = \max_{i=1,\ldots,n} X_i$ satisfies $\Pr(M_n \le x) = \{\Pr(X_1 \le x)\}^n$. Although this is no longer true in the presence of dependence, certain mixing conditions still guarantee that $\Pr(M_r \le x)$ is close to $\{\Pr(M_m \le x)\}^{r/m}$ for suitable r and m. This is important in so far it implies that for a broad class of stationary sequences the only non-trivial weak limits of scaled and normalized sample maxima are the extreme-value distributions, whose range of applicability is thereby greatly enlarged (Leadbetter *et al.* 1983). The argument can be extended to the multivariate case (Hsing 1989).

A natural question, then, is whether in general the probability q_r of no extreme event in a block of size r can be approximated by the probability $q_m^{r/m}$ of no extreme event in r/m independent smaller blocks of size m. Finite-sample inequalities in Subsection 3.1 lead to asymptotic results in Subsection 3.2.

3.1 Inequalities

Let $A_1, \ldots, A_r$ be block-stationary events (Definition 2.1) and employ the notations of Section 2. Two lemmas will prepare the ground for inequalities for $q_r - q_m^{r/m}$ in case m is small compared to r (Theorem 3.4) and inequalities for $q_r - q_s^{r/s}$ in case s can be of the same order as r (Theorem 3.5). By convention, we set the sum over the empty set equal to zero and the product over the empty set equal to one.

Lemma 3.1 *Let $a_1, b_1, \ldots, a_k, b_k \in \{0, \ldots, r\}$ and assume that there exists a positive integer l such that $b_i - a_i \ge l$ for all $i = 1, \ldots, k$ and $a_{i+1} - b_i = l$ for all $i = 1, \ldots, k-1$.*

We have

$$-(\alpha_l + p_l)\sum_{i=2}^{k}\prod_{j=i+1}^{k} q_{b_j-a_j} \le q_{b_k-a_1} - \prod_{i=1}^{k} q_{b_i-a_i} \le \alpha_l \sum_{i=2}^{k}\prod_{j=i+1}^{k} q_{b_j-a_j}.$$

Proof. We proceed by induction on k. For $k = 1$, there is nothing to prove. Let $k \ge 2$. We have

$$\begin{aligned}&\Pr\left(\bigcap_{i=a_1+1}^{b_{k-1}} A_i^c \cap \bigcap_{i=a_k+1}^{b_k} A_i^c\right) - \Pr\left(\bigcup_{i=b_{k-1}+1}^{a_k} A_i\right)\\ \le\ &\Pr\left(\bigcap_{i=a_1+1}^{b_k} A_i^c\right) = q_{b_k-a_1} \le \Pr\left(\bigcap_{i=a_1+1}^{b_{k-1}} A_i^c \cap \bigcap_{i=a_k+1}^{b_k} A_i^c\right).\end{aligned}$$

Moreover,

$$\left|\Pr\left(\bigcap_{i=a_1+1}^{b_{k-1}} A_i^c \cap \bigcap_{i=a_k+1}^{b_k} A_i^c\right) - q_{b_{k-1}-a_1} q_{b_k-a_k}\right| \le \alpha_l.$$

Together, we find

$$q_{b_{k-1}-a_1} q_{b_k-a_k} - \alpha_l - p_l \le q_{b_k-a_1} \le q_{b_{k-1}-a_1} q_{b_k-a_k} + \alpha_l.$$

Apply the induction hypothesis on $q_{b_{k-1}-a_1}$ to conclude the proof. □

For a real number x, we denote by $\lfloor x \rfloor$ the largest integer not larger than x, and by $\lceil x \rceil$ the smallest integer not smaller than x.

Lemma 3.2 *Let l and m be positive integers such that $l \le m \le r$. For every $k = 1, \dots, \lfloor (r+l)/(m+l) \rfloor$, we have*

$$q_r \le q_m^k + \alpha_l \frac{1-q_m^{k-1}}{1-q_m}.$$

If also $2l + m \le r$, then for $k = \lceil (r+l)/(m+l) \rceil$, we have

$$q_r \ge q_m^k - (\alpha_l + p_l)\frac{1-q_m^{k-1}}{1-q_m}.$$

Proof. Let $k = 1, \dots, \lfloor (r+l)/(m+l) \rfloor$ and set $a_i = (i-1)(m+l)$ and $b_i = a_i + m$ for $i = 1, \dots, k$. The integers $a_1, b_1, \dots, a_k, b_k$ satisfy the conditions of Lemma 3.1; in particular $b_k = km + (k-1)l \le r$. Hence

$$-(\alpha_l + p_l)\sum_{i=2}^{k} q_m^{k-i} \le q_{km+(k-1)l} - q_m^k \le \alpha_l \sum_{i=2}^{k} q_m^{k-i}.$$

Now we have $\sum_{i=2}^{k} q_m^{k-i} = (1-q_m^{k-1})/(1-q_m)$. Since $q_r \leq q_{km+(k-1)l}$, the upper bound follows.

Next, suppose that $2l+m \leq r$. Apply Lemma 3.1 on $a_1 = 0$, $b_1 = m$, $a_2 = m+l$, and $b = r$ to find

$$q_r \geq q_m q_{r-m-l} - (\alpha_l + p_l).$$

Let $k = \lceil (r+l)/(m+l) \rceil$. Since $r-m-l \leq (k-1)(m+l) - l \leq r$, we have

$$q_{r-m-l} \geq q_{(k-1)(m+l)-l} \geq q_m^{k-1} - (\alpha_l + p_l)\frac{1-q_m^{k-2}}{1-q_m}.$$

Substitute the lower bound for q_{r-m-l} into the lower bound for q_r to conclude the proof. □

Remark 3.3 For $0 < x < 1$ and $a \geq 1$ or $a = 0$, we have

$$\frac{1-x^a}{1-x} \leq \min(a, 1/(1-x)).$$

Hence, $(1-q_m^{k-1})/(1-q_m) \leq \min(k-1, 1/p_m)$ in Lemma 3.2.

Theorem 3.4 *For positive integers l and m such that $l \leq m \leq r$, we have*

$$q_r \leq q_m^{r/m} + \alpha_l \frac{1-q_m^{r/m}}{1-q_m} + \frac{l}{m} + \frac{m}{r}.$$

If also $2l+m \leq r$, then

$$q_r \geq q_m^{r/m} - (\alpha_l + p_l)\frac{1-q_m^{r/m}}{1-q_m} - \frac{l}{m} - \frac{m}{r}.$$

Proof. Lemma 3.2 with $k = \lfloor (r+l)/(m+l) \rfloor$ gives

$$q_r \leq q_m^k + \alpha_l \frac{1-q_m^{k-1}}{1-q_m}.$$

If $0 \leq x \leq 1$, $a > 0$, and $b > 0$, then $|x^a - x^b| \leq \max(1-a/b, 1-b/a)$. Hence, since $(r+l)/(m+l) - 1 \leq k \leq (r+l)/(m+l) \leq r/m$, we have

$$q_m^k - q_m^{r/m} \leq 1 - mk/r \leq l/m + m/r,$$

leading to the stated upper bound for q_r.

Next, suppose $2l+m \leq r$, and set $k = \lceil (r+l)/(m+l) \rceil$. By Lemma 3.2, we have

$$q_r \geq q_m^k - (\alpha_l + p_l)\frac{1-q_m^{k-1}}{1-q_m}.$$

Now, by the same inequality as before, we have

$$|q_m^k - q_m^{r/m}| \leq \max(1-mk/r, 1-r/(mk)).$$

Since $(r+l)/(m+l) \leq k < (r+l)/(m+l)+1$, we have $1-mk/r \leq l/m$ and $1-r/(mk) \leq (l+m)/r$, so that $\max(1-mk/r, 1-r/(mk)) \leq l/m + m/r$. As $k-1 \leq r/m$, the proof is complete. □

Theorem 3.5 *For positive integers l, m, and s such that $l \le m$, $m + 2l \le s$, and $s \le r$, we have*

$$\left| q_r - q_s^{r/s} \right| \le \frac{r}{m}(2\alpha_l + p_l) + 4\frac{m}{s}.$$

Proof. Let $k = \lfloor (s+l)/(m+l) \rfloor$. By Lemma 3.2, we have

$$q_s \le q_m^k + (k-1)\alpha_l.$$

Since $\{\min(x+y, 1)\}^a \le x^a + ay$ for $0 \le x \le 1$, $y \ge 0$, and $a \ge 1$, we have

$$q_s^{r/s} \le q_m^{kr/s} + \frac{r}{s}(k-1)\alpha_l.$$

Let $j = \lceil (r+l)/(m+l) \rceil$. By Lemma 3.2, we also have

$$q_r \ge q_m^j - (j-1)(\alpha_l + p_l).$$

All in all, we find

$$q_r - q_s^{r/s} \ge q_m^j - q_m^{kr/s} - \left(\frac{r}{s}(k-1) + (j-1) \right) \alpha_l - (j-1)p_l.$$

Now if $j \le kr/s$, then $q_m^j - q_m^{kr/s} \ge 0$, while if $j > kr/s$, then

$$q_m^j - q_m^{kr/s} \ge \frac{kr}{js} - 1 \ge \frac{(s-m)r}{(r+2l+m)s} - 1 \ge -4\frac{m}{s}.$$

Further, $j - 1 < (r+l)/(m+l) < r/m$, and $(r/s)(k-1) < r/m$.

The proof of the upper bound for $q_r - q_s^{r/s}$ is analogous, and based on the inequality $\{\max(x-y, 0)\}^a \ge x^a - ay$ for $0 \le x \le 1$, $y \ge 0$, and $a \ge 1$. □

3.2 Asymptotic results

For every $n \ge 1$ let $A_{1,n}, \dots, A_{r_n,n}$ be block-stationary events, and use the notations of Section 2. Theorem 3.7 compares $q_{r_n,n}$ with $q_{s_n,n}^{s_n/r_n}$.

Lemma 3.6 *Let $1 \le l_n \le m_n \le r_n$ be integers with $l_n = o(m_n)$.*

(i) Let $0 < \lambda_n \to 0$. If $p_{m_n,n} = O(\lambda_n)$ and $\alpha_{l_n,n} = o(\lambda_n)$, then $p_{l_n,n} = o(\lambda_n)$.

(ii) If $0 < p_{m_n,n} \to 0$ and $\alpha_{l_n,n} = o(p_{m_n,n})$, then $p_{l_n,n} = o(p_{m_n,n})$.

Proof. (i) Let k be a positive integer. If n is large enough so that $k \le \lfloor (m_n + l_n)/(2l_n) \rfloor$, then we have by Lemma 3.2,

$$1 - p_{m_n,n} \le (1 - p_{l_n,n})^k + (k-1)\alpha_{l_n,n} \le \exp(-p_{l_n,n}k) + (k-1)\alpha_{l_n,n}.$$

If n is also large enough so that $p_{m_n,n} + (k-1)\alpha_{l_n,n} < 1$, then

$$p_{l_n,n} \le -\frac{1}{k} \log\{1 - p_{m_n,n} - (k-1)\alpha_{l_n,n}\}.$$

Hence we have

$$\limsup_{n\to\infty} p_{l_n,n}/\lambda_n \le \frac{1}{k} \limsup_{n\to\infty} p_{m_n,n}/\lambda_n.$$

Let $k \to \infty$ to see that $p_{l_n,n}/\lambda_n \to 0$.

(ii) Take $\lambda_n = p_{m_n,n}$ in (i). $\square$

Theorem 3.7 *For positive integers l_n and s_n such that*

$$l_n = o(s_n), \quad s_n \le r_n, \quad \textit{and} \quad \alpha_{l_n,n} = o(\max(s_n/r_n, p_{s_n,n})),$$

we have

$$q_{r_n,n} = q_{s_n,n}^{r_n/s_n} + o(1).$$

If additionally $\liminf_{n\to\infty} s_n/r_n > 0$*, then also*

$$q_{r_n,n}^{s_n/r_n} = q_{s_n,n} + o(1).$$

Proof. Without loss of generality, we can restrict n to a subsequence along which $(r_n/s_n)p_{s_n,n} \to c \in [0,\infty]$.

If $c = \infty$, then $q_{s_n,n}^{r_n/s_n} \to 0$. Set $k_n = \lfloor (r_n + l_n)/(s_n + l_n) \rfloor$. By Lemma 3.2, we have

$$q_{r_n} \le q_{s_n,n}^{k_n} + \alpha_{l_n,n}/p_{s_n,n} \to 0,$$

since $k_n \sim r_n/s_n$.

Next, suppose $c < \infty$. Since in this case $\alpha_{l_n,n} = o(s_n/r_n)$, we can find positive integers m_n such that

$$l_n = o(m_n), \quad m_n = o(s_n), \quad \text{and} \quad \alpha_{l_n,n} = o(m_n/r_n).$$

Again, without loss of generality, we can restrict n to a further subsequence such that $(r_n/m_n)p_{m_n} \to d \in [0,\infty]$.

Suppose first that $d < \infty$. By Lemma 3.6(i), we have $(r_n/m_n)p_{l_n,n} \to 0$. But then $|q_{r_n,n} - q_{s_n,n}^{r_n/s_n}| \to 0$ by Theorem 3.5.

Suppose next that $d = \infty$. Let $j_n = \lfloor (r_n + l_n)/(m_n + l_n) \rfloor$. Since $j_n \sim r_n/m_n$, we have by Lemma 3.2,

$$q_{r_n,n} \le (1 - p_{m_n,n})^{j_n} + (j_n - 1)\alpha_{l_n,n} \to 0.$$

Next, let $k_n = \lfloor (s_n + l_n)/(m_n + l_n) \rfloor$. By Lemma 3.2, we have

$$q_{s_n,n}^{r_n/s_n} \le \min\{(1 - p_{m_n,n})^{k_n} + (k_n - 1)\alpha_{l_n,n}, 1\}^{r_n/s_n}.$$

If $a \ge 1$, $0 \le x \le 1$, and $y \ge 0$, then $\min(x + y, 1)^a \le x^a + ay$. Hence

$$q_{s_n,n}^{r_n/s_n} \le (1 - p_{m_n,n})^{k_n r_n/s_n} + (r_n/s_n)(k_n - 1)\alpha_{l_n,n}.$$

Since $k_n \sim s_n/m_n$, we obtain $q_{s_n,n}^{r_n/s_n} \to 0$.

The second statement of the Theorem follows from

$$q_{s_n,n} = \{\max(q_{r_n,n} + o(1), 0)\}^{s_n/r_n}$$

and the uniform continuity of the map $(x, a) \mapsto x^a$ on $(x, a) \in [0, 2] \times [\epsilon, 1]$, where $0 < \epsilon \leq 1$. □

4 Application: disjoint or sliding blocks?

Estimation of the distribution of the maximum of a block of consecutive variables lies at the heart of the method of annual maxima (Gumbel 1958) and the blocks estimator for the extremal index (Hsing 1991; Smith and Weissman 1994). In each case a sample of observations is partitioned into blocks to yield a sample of block maxima, from which the unknown distribution can be estimated. An alternative to disjoint blocks is to slide a window of the appropriate size through the sample. The resulting sample of sliding-block maxima is much larger than the one from disjoint blocks; however, block maxima of overlapping windows are dependent, even in case of independent observations. This raises the question which are the more efficient: disjoint or sliding blocks?

The problem can be solved in our general framework. For every n, let $A_{1,n}, \ldots, A_{n,n}$ be block-stationary events on a common probability space (which may vary with n). For $k = 1, \ldots, n$ and $j = 0, \ldots, k-1$, let $I_{j,k,n}$ be the indicator function of the event $\bigcap_{i=j+1}^{k} A_{i,n}^c$. Abbreviate $I_{k,n} = I_{0,k,n}$ for $k = 1, \ldots, n$. Observe that $q_{r,n} = E(I_{r,n}) = E(I_{j,j+r,n})$, for $r = 1, \ldots, n$ and $j = 0, \ldots, n-r$. We can express the familiar mixing coefficients by

$$\begin{aligned} \alpha_{s,l,n} \;&=\; \max\{|\mathrm{Cov}\,(I_{u,v,n}, I_{s+v,s+w,n})\,| : \\ &\qquad u \geq 0, v - u \geq l, w - v \geq l, w + s \leq n\}. \end{aligned}$$

Put $\bar{\alpha}_{l,n} = \max_{s=l,\ldots,n} \alpha_{s,l,n}$. Two unbiased estimators of $q_{r,n}$ are

$$\hat{q}_{r,n} = \frac{1}{\lfloor n/r \rfloor} \sum_{i=1}^{\lfloor n/r \rfloor} I_{(i-1)r,ir,n} \quad \text{and} \quad \tilde{q}_{r,n} = \frac{1}{n - r + 1} \sum_{i=0}^{n-r} I_{i,i+r,n},$$

composed of disjoint blocks and sliding blocks respectively. By Theorem 4.1 below and the inequality $2x\{(\log x)^{-1}(x-1) - x\} < x(1-x)$ for $0 < x < 1$, the sliding-blocks estimator is more efficient than the disjoint-blocks estimator in case $q_{r,n}$ is bounded away from 0 and 1.

Theorem 4.1 *If the positive integers l_n and r_n are such that*

$$l_n = o(r_n), \quad r_n = o(n), \quad \textit{and} \quad \bar{\alpha}_{l_n,n} = o(r_n/n),$$

then

$$\begin{aligned} (n/r_n)\mathrm{Var}\,(\hat{q}_{r_n,n}) \;&=\; q_{r_n,n}(1 - q_{r_n,n}) + o(1), \\ (n/r_n)\mathrm{Var}\,(\tilde{q}_{r_n,n}) \;&=\; 2q_{r_n,n}\left(\frac{q_{r_n,n} - 1}{\log(q_{r_n,n})} - q_{r_n,n}\right) + o(1). \end{aligned}$$

Proof. We give the proof only for the sliding-blocks estimator, which is the more difficult part. We have

$$\operatorname{Var}(\tilde{q}_{r_n,n}) = \frac{1}{(n-r_n+1)^2} \sum_{i=0}^{n-r_n} \sum_{j=0}^{n-r_n} \operatorname{Cov}(I_{i,i+r_n,n}, I_{j,j+r_n,n})$$

Since

$$\begin{aligned} \operatorname{Cov}(I_{i,i+r_n,n}, I_{j,j+r_n,n}) &= q_{r_n+|i-j|,n} - q_{r_n,n}^2 && \text{if } |i-j| \le r_n, \\ |\operatorname{Cov}(I_{i,i+r_n,n}, I_{j,j+r_n,n})| &\le \bar{\alpha}_{l_n,n} && \text{if } |i-j| \ge r_n + l_n, \end{aligned}$$

we have

$$\begin{aligned} \operatorname{Var}(\tilde{q}_{r_n,n}) &= \frac{q_{r_n,n}(1-q_{r_n,n})}{n-r_n+1} + \frac{2}{(n-r_n+1)^2} \sum_{h=1}^{r_n} (n-r_n+1-h)(q_{r_n+h,n} - q_{r_n,n}^2) \\ &\quad + O(l_n/n) + O(\bar{\alpha}_{l_n,n}) \\ &= \frac{2}{n-r_n+1} \sum_{h=1}^{r_n} (q_{r_n+h,n} - q_{r_n,n}^2) + O(r_n^2/n^2) + O(l_n/n) + O(\bar{\alpha}_{l_n,n}). \end{aligned}$$

We obtain

$$\begin{aligned} (n/r_n)\operatorname{Var}(\tilde{q}_{r_n,n}) &= \frac{2}{r_n} \sum_{h=1}^{r_n} (q_{r_n+h,n} - q_{r_n,n}^2) + o(1) \\ &= 2 \int_0^1 (q_{r_n+\lceil r_n x \rceil,n} - q_{r_n,n}^2) \, \mathrm{d}x + o(1). \end{aligned}$$

By Theorem 3.7, we have $q_{r_n+\lceil r_n x \rceil,n} = q_{r_n,n}^{1+\lceil r_n x \rceil/r_n} + o(1) = q_{r_n,n}^{1+x} + o(1)$ for $x \ge 0$. Apply the dominated convergence theorem to complete the proof. □

5 After an extreme event

The extremal index θ of a stationary sequence $\{X_n : n \ge 1\}$ of real-valued random variables determines the dependence between extreme events in the sequence in a number of different ways. The results to follow are nothing but generalizations of these facts, under minimal conditions, to the naked framework of block-stationary events.

Let us first recall some properties of the extremal index. Denote the marginal distribution function of the X_n by F, and for $n \ge 1$ let the thresholds $u_n \in \mathbb{R}$ be such that $\limsup n[1 - F(u_n)] < \infty$. Denote

$$\theta_{m,n} = \Pr\left(\max_{i=2,\ldots,m} X_i \le u_n \,\middle|\, X_1 > u_n \right),$$

the conditional probability that a threshold exceedance is followed by a run of non-exceedances. O'Brien (1987) proved that if $m \equiv m_n \to \infty$ and $m_n = o(n)$, then, under certain mixing conditions, $\Pr(M_n \le u_n) = \{F(u_n)\}^{n\theta_{m_n,n}} + o(1)$. Hence, the extremal index arises as the limit of $\theta_{m_n,n}$.

Another characterization of the extremal index is in terms of $S_{m,n} = \sum_{i=1}^m I(X_i > u_n)$, the number of threshold exceedances in a block of size m. If $m \equiv m_n = o(n)$ and $\limsup n[1 - F(u_n)] < \infty$, then $\Pr(S_{m_n,n} > 0) \le m_n[1 - F(u_n)] \to 0$. In case $S_{m,n} > 0$ all the exceedances in the block are thought of as one single cluster. If $m_n \to \infty$, then under certain mixing conditions the expected cluster size satisfies $E[S_{m_n,n} \mid S_{m_n,n} > 0] \to 1/\theta$ (Leadbetter 1983). In words, the extremal index is the reciprocal of the mean cluster size of threshold exceedances.

Finally, Ferro and Segers (2002) linked the extremal index with the inter-exceedance times $T(u_n) \stackrel{d}{=} \min\{i \ge 1 : X_{i+1} > u_n\}$ conditionally on $X_1 > u_n$. They showed that, again under certain mixing conditions,

$$\Pr\{[1 - F(u_n)]T(u_n) > x \mid X_1 > u_n\} \to \theta \exp(-x\theta), \quad \text{for } x > 0,$$

that is, the normalized inter-exceedance times $[1 - F(u_n)]T(u_n)$ converge to a mixture between a point mass at zero and the exponential distribution.

The proper reformulations of these properties in terms of the threshold exceedances $\{X_i > u_n\}$ will be shown to remain true in the general setting of row-wise block-stationary events $A_{i,n}$. The asymptotic results of Subsection 5.2 are founded on the finite-sample inequalities of Subsection 5.1 and culminate in the Characterization Theorem of Subsection 5.3.

5.1 Inequalities

Let $A_1, \ldots, A_r$ be block-stationary events (Definition 2.1) and recall the notations of Section 2.

Theorem 5.1 *For $m = 1, \ldots, r$, we have*

$$mp\theta_m \le p_m.$$

If, additionally, $2m < r$ and $l = 1, \ldots, \min(m, r - 2m)$, then also

$$p_m \le mp\theta_m + p_m^2 + p_l + \alpha_l.$$

Proof. Since $\theta_i \ge \theta_{i+1}$ for all $i = 1, \ldots, r - 1$, we have

$$p_m = \sum_{i=1}^{m}(p_i - p_{i-1}) = \sum_{i=1}^{m} p\theta_i \ge mp\theta_m.$$

For the upper bound, observe that

$$p_m = \Pr\left(\bigcup_{i=1}^{m} A_i \cap \bigcap_{i=m+1}^{2m+l} A_i^c\right) + \Pr\left(\bigcup_{i=1}^{m} A_i \cap \bigcup_{i=m+1}^{2m+l} A_i\right).$$

On the one hand, we have

$$\Pr\left(\bigcup_{i=1}^{m} A_i \cap \bigcap_{i=m+1}^{2m+l} A_i^c\right) = \sum_{i=1}^{m} \Pr\left(A_i \cap \bigcap_{j=i+1}^{2m+l} A_j^c\right) = \sum_{i=1}^{m} p\theta_{2m+l-i+1} \leq mp\theta_m,$$

while on the other hand, as $\Pr(A^c \cap B) - \Pr(A^c)\Pr(B) = \Pr(A)\Pr(B) - \Pr(A \cap B)$ for general events A and B, we have

$$\Pr\left(\bigcup_{i=1}^{m} A_i \cap \bigcup_{i=m+1}^{2m+l} A_i\right) \leq \Pr\left(\bigcup_{i=1}^{m} A_i \cap \bigcup_{i=m+l+1}^{2m+l} A_i\right) + p_l \leq p_m^2 + \alpha_l + p_l.$$

□

Theorem 5.2 *For positive integers l and m such that $l \leq m \leq r$, we have*

$$q_r \leq (1-\theta_m p)^r + \alpha_l \frac{1-q_m^{r/m}}{1-q_m} + \frac{l}{m} + \frac{m}{r}.$$

If additionally $2m + l \leq r$, then

$$q_r \geq \exp(-r\theta_m p) - \frac{2r}{m}(\alpha_l + p_l) - \frac{r}{m}p_m^2 - \frac{l}{m} - 2\frac{m}{r}.$$

Proof. By Theorem 3.4, we have

$$q_r \leq q_m^{r/m} + \alpha_l \frac{1-q_m^{r/m}}{1-q_m} + \frac{l}{m} + \frac{m}{r}.$$

By the first inequality of Theorem 5.1, we have $q_m \leq 1 - mp\theta_m$. Since $(1+x)^a \leq 1 + ax$ for $x \geq -1$ and $0 < a \leq 1$, we obtain

$$q_m^{r/m} \leq (1 - mp\theta_m)^{r/m} \leq (1-\theta_m p)^r,$$

which gives the stated upper bound for q_r.

Secondly, we have by Theorem 3.4,

$$q_r \geq q_m^{r/m} - \frac{r}{m}(\alpha_l + p_l) - \frac{l}{m} - \frac{m}{r}.$$

The second inequality of Theorem 5.1 implies

$$q_m^{r/m} = (1-p_m)^{r/m} \geq \{\max(1 - mp\theta_m - p_m^2 - p_l - \alpha_l, 0)\}^{r/m}.$$

Since $\{\max(1-x, 0)\}^a \geq \exp(-ax) - 1/a$ for $x \geq 0$ and $a > 0$, we have

$$q_m^{r/m} \geq \exp\{-rp\theta_m - (r/m)(p_m^2 + p_l + \alpha_l)\} - m/r.$$

Since also $\exp(-x-y) \geq \exp(-x) - y$ for $x \geq 0$ and $y \geq 0$, we get

$$q_m^{r/m} \geq \exp(-rp\theta_m) - (r/m)(p_m^2 + p_l + \alpha_l) - m/r.$$

Substitute this inequality in the lower bound for q_r to conclude the proof. □

Theorem 5.3 *For positive integers l and m such that $l \leq m$ and $2m + l \leq r$, and for $s = m + l, \ldots, r - m$, we have*

$$\begin{aligned} -(mp)^{-1}\alpha_l - (mp)^{-1}p_l &\leq \theta_s - \theta_m q_s \\ &\leq 3(mp)^{-1}\bar{\alpha}_l + 2p_m + (1 + (mp)^{-1})p_l. \end{aligned}$$

Proof. For $t = m + 1, \ldots, r$, we have

$$\Pr\left(\bigcup_{i=1}^{m} A_i \cap \bigcap_{i=m+1}^{t} A_i^c\right) = \sum_{k=1}^{m} \Pr\left(A_k \cap \bigcup_{i=k+1}^{t} A_i^c\right) = \sum_{k=1}^{m} p\theta_{t-k+1},$$

so that

$$mp\theta_t \leq \Pr\left(\bigcup_{i=1}^{m} A_i \cap \bigcap_{i=m+1}^{t} A_i^c\right) \leq mp\theta_{t-m}.$$

Hence for $s = m + 1, \ldots, r - m$, we have

$$\Pr\left(\bigcup_{i=1}^{m} A_i \cap \bigcap_{i=m+1}^{s+m} A_i^c\right) \leq mp\theta_s \leq \Pr\left(\bigcup_{i=1}^{m} A_i \cap \bigcap_{i=m+1}^{s} A_i^c\right).$$

Now

$$\begin{aligned} 0 &\leq \Pr\left(\bigcup_{i=1}^{m} A_i \cap \bigcap_{i=m+1}^{s} A_i^c\right) - \Pr\left(\bigcup_{i=1}^{m} A_i \cap \bigcap_{i=m+1}^{s+m} A_i^c\right) \\ &\leq \Pr\left(\bigcup_{i=1}^{m} A_i \cap \bigcup_{i=s+1}^{s+m} A_i\right) \leq p_m^2 + \alpha_{s-m,l}. \end{aligned}$$

Moreover,

$$0 \leq \Pr\left(\bigcup_{i=1}^{m} A_i \cap \bigcap_{i=m+l+1}^{s+m} A_i^c\right) - \Pr\left(\bigcup_{i=1}^{m} A_i \cap \bigcap_{i=m+1}^{s+m} A_i^c\right) \leq \Pr\left(\bigcup_{i=m+1}^{m+l} A_i\right) = p_l$$

and, if $s \geq m + l$,

$$\left|\Pr\left(\bigcup_{i=1}^{m} A_i \cap \bigcap_{i=m+l+1}^{s+m} A_i^c\right) - p_m q_{s-l}\right| \leq \alpha_l.$$

Together, we obtain

$$p_m q_{s-l} - \alpha_l - p_l \leq mp\theta_s \leq p_m q_{s-l} + \alpha_l + p_m^2 + \alpha_{s-m,l}.$$

If $s - m \geq l$, then $\alpha_{s-m,l} \leq \bar{\alpha}_l$. Theorem 5.1 now implies

$$mp\theta_m q_{s-l} - \alpha_l - p_l \leq mp\theta_s \leq mp\theta_m q_{s-l} + 3\bar{\alpha}_l + 2p_m^2 + p_l.$$

Since $q_s \leq q_{s-l} \leq q_s + p_l$, we obtain

$$mp\theta_m q_s - \alpha_l - p_l \leq mp\theta_s \leq mp\theta_m q_s + 3\bar{\alpha}_l + 2p_m^2 + (1 + mp)p_l.$$

Divide by mp and use $p_m \leq mp$ to conclude the proof. □

5.2 Asymptotic results

For every $n \geq 1$, let the events $A_{1,n}, \ldots, A_{r_n,n}$ be block-stationary. Recall the notations of the Introduction, in particular

$$\theta_{m,n} = \Pr\left(\bigcap_{i=2}^{m} A_{i,n}^c \,\middle|\, A_{1,n} \right) = (p_{m,n} - p_{m-1,n})/p_n, \quad \text{for } m = 2, \ldots, r_n.$$

Consecutively treated in this paragraph are extremes in small blocks, extremes in large blocks, and inter-arrival times between extreme events.

Small blocks

Theorem 5.4 *Let l_n and m_n be positive integers with $2m_n + l_n \leq r_n$ and assume that $l_n = o(m_n)$ and $p_{m_n,n} \to 0$.*

(i) If $\alpha_{l_n} = o(m_n p_n)$, then $\theta_{m_n,n} = (m_n p_n)^{-1} p_{m_n,n} + o(1)$.

(ii) If $\alpha_{l_n,n} = o(p_{m_n,n})$, then $\theta_{m_n,n} \sim (m_n p_n)^{-1} p_{m_n,n}$.

Proof. (i) By Theorem 5.1, we have

$$(m_n p_n)^{-1} \left(p_{m_n,n} - p_{m_n,n}^2 - p_{l_n,n} - \alpha_{l_n,n} \right) \leq \theta_{m_n,n} \leq (m_n p_n)^{-1} p_{m_n,n}.$$

Since $p_{m_n,n} \leq m_n p_n$ and $p_{l_n,n} \leq l_n p_n$, statement (i) follows.

(ii) We can also write the previously displayed inequalities as

$$\frac{p_{m_n,n}}{m_n p_n} \left(1 - p_{m_n,n} - \frac{p_{l_n,n}}{p_{m_n,n}} - \frac{\alpha_{l_n,n}}{p_{m_n,n}} \right) \leq \theta_{m_n,n} \leq (m_n p_n)^{-1} p_{m_n,n}.$$

By Lemma 3.6(ii), we have $p_{l_n,n} = o(p_{m_n,n})$, hence (ii) follows. □

Remark 5.5 Theorem 5.4 has the following interpretation. For $i = 1, \ldots, r_n$, let $I_{i,n}$ be the indicator of the event $A_{i,n}$, and for $m = 1, \ldots, r_n$, let $S_{m,n} = \sum_{i=1}^m I_{i,n}$ be the number of extreme events that occurred in a block of size m. Then $E(S_{m,n} \mid S_{m,n} > 0) = mp_n/p_{m,n}$ for $m = 1, \ldots, r$, so that under the conditions of Theorem 5.4(i), we may write

$$\theta_{m_n,n} = [E(S_{m_n,n} \mid S_{m_n,n} > 0)]^{-1} + o(1).$$

Under the conditions of Theorem 5.4(ii), we also have

$$E(S_{m_n,n} \mid S_{m_n,n} > 0) = \theta_{m_n,n}^{-1} + o(1).$$

Theorem 5.6 *Let l_n, m_n, and M_n be positive integers such that $l_n \leq m_n \leq M_n$ and $2M_n + l_n \leq r_n$. If*

$$l_n = o(m_n), \quad \alpha_{l_n,n} = o(m_n p_n), \quad p_{M_n,n} \to 0, \quad \textit{and} \quad M_n p_n = O(1),$$

then

$$\theta_{m_n,n} = \frac{p_{m_n,n}}{m_n p_n} + o(1) = \frac{p_{M_n,n}}{M_n p_n} + o(1) = \theta_{M_n,n} + o(1).$$

Proof. By Theorem 5.4, we have immediately that

$$\theta_{m_n,n} = \frac{p_{m_n,n}}{m_n p_n} + o(1) \quad \text{and} \quad \theta_{M_n,n} = \frac{p_{M_n,n}}{M_n p_n} + o(1).$$

So it is enough to show that the two right-hand sides of these equations are equal.

Suppose first that $m_n = o(M_n)$. On the one hand, we have

$$p_{M_n,n} \leq p_{m_n \lceil M_n/m_n \rceil, n} \leq \lceil M_n/m_n \rceil p_{m_n,n} = (M_n/m_n) p_{m_n,n}[1 + o(1)],$$

and thus $(M_n p_n)^{-1} p_{M_n,n} \leq (m_n p_n)^{-1} p_{m_n,n} + o(1)$. On the other hand, we have by Lemma 3.2, with $k_n = \lfloor (M_n + l_n)/(m_n + l_n) \rfloor$,

$$1 - p_{M_n,n} - (k_n - 1)\alpha_{l_n,n} \leq (1 - p_{m_n,n})^{k_n} \leq \exp(-k_n p_{m_n,n}).$$

Since $k_n \leq M_n/m_n = O(1/(m_n p_n))$, we can take logarithms of both sides of the displayed equation. We find

$$k_n p_{m_n,n} \leq -\log\{1 - p_{M_n,n} - (k_n - 1)\alpha_{l_n,n}\},$$

and hence

$$(M_n/m_n) p_{m_n,n} \leq (p_{M_n,n} + k_n \alpha_{l_n,n})[1 + o(1)].$$

Since $k_n/M_n \sim 1/m_n$, we find $(m_n p_n)^{-1} p_{m_n,n} \leq (M_n p_n)^{-1} p_{M_n,n} + o(1)$.

Next consider the general case $m_n \leq M_n$. We can find positive integers m_n' such that $l_n = o(m_n')$ and $\alpha_{l_n,n} = o(m_n' p_n)$. By the previous argument, we have

$$\frac{p_{m_n',n}}{m_n' p_n} = \frac{p_{m_n,n}}{m_n p_n} + o(1) = \frac{p_{M_n,n}}{M_n p_n} + o(1).$$

□

Big blocks

Theorem 5.7 *Let l_n and m_n be positive integers. If*

$$l_n = o(m_n), \quad m_n = o(r_n), \quad \text{and} \quad \alpha_{l_n,n} = o(\max(m_n/r_n), p_{m_n,n}),$$

then

$$q_{r_n,n} \leq (1 - \theta_{m_n,n} p_n)^{r_n} + o(1) = \exp(-r_n \theta_{m_n,n} p_n) + o(1).$$

If additionally $p_{m_n,n} \to 0$, then

$$\begin{aligned} q_{r_n,n} &= (1 - \theta_{m_n,n} p_n)^{r_n} + o(1) = \exp(-r_n \theta_{m_n,n} p_n) + o(1) \\ &= (1 - p_n)^{r_n \theta_{m_n,n}} + o(1). \end{aligned}$$

Proof. By the first inequality of Theorem 5.2, we immediately have

$$q_{r_n,n} \leq (1-\theta_{m_n,n}p_n)^{r_n} + o(1).$$

Since $0 \leq \exp(-ax) - (1-x)^a \leq 1/a$ for $0 \leq x \leq 1$ and $a > 0$, we also have

$$(1-\theta_{m_n,n}p_n)^{r_n} = \exp(-r_n\theta_{m_n,n}p_n) + O(1/r_n).$$

Next suppose $p_{m_n,n} \to 0$. As also $p_n \leq p_{m_n,n} \to 0$, we have

$$(1-p_n)^{r_n\theta_{m_n,n}} = \exp(-r_n\theta_{m_n,n}p_n) + o(1),$$

since $\sup_{a\geq 0} |\exp(-ax)-(1-x)^a| \to 0$ as $0 < x \to 0$. So we only need to prove that $q_{r_n,n} \geq \exp(-r_n\theta_{m_n,n}p_n) + o(1)$. Without loss of generality, we may restrict n to a subsequence along which $(r_n/m_n)p_{m_n,n}$ converges to some $c \in [0,\infty]$.

If $c < \infty$, then $(r_n/m_n)p_{l_n,n} \to 0$ by Lemma 3.6(i). The second inequality of Theorem 5.2 now finishes the job.

If $c = \infty$, then $m_n/r_n = o(p_{m_n,n})$, so that $\alpha_{l_n,n} = o(p_{m_n,n})$. By Theorem 5.4, we have

$$r_n\theta_{m_n,n}p_n \sim (r_n/m_n)p_{m_n,n} \to \infty,$$

so that $\exp(-r_n\theta_{m_n,n}p_n) \to 0$. □

Remark 5.8 Without the extra condition $p_{m_n,n} \to 0$, the second statement of Theorem 5.7 is not true. Consider for example independent events with $p_n \to 0$, $r_n \sim p_n^{-3}$, and $m_n \sim p_n^{-2}$: we have $q_{r_n,n} = (1-p_n)^{r_n} \to 0$, but $r_n\theta_{m_n,n}p_n \sim p_n^{-2}(1-p_n)^{m_n-1} \to 0$.

The condition $p_{m_n,n} \to 0$ is implied by each of the following ones: (i) $m_np_n \to 0$, (ii) $\limsup_{n\to\infty} r_np_n < \infty$, and (iii) $\liminf_{n\to\infty} q_{r_n,n} > 0$. Regarding (i), just observe that $p_{m_n,n} \leq m_np_n$. Since $m_n = o(r_n)$, (ii) implies (i). And since $q_{r_n,n} = (1-p_{m_n,n})^{r_n/m_n} + o(1)$ by Theorem 3.7, also condition (iii) is sufficient.

Inter-arrival times between extreme events

Theorem 5.9 *Let l_n and m_n be positive integers such that $2m_n + l_n \leq r_n$. If*

$$l_n = o(m_n), \quad m_np_n \to 0, \quad \text{and} \quad \bar{\alpha}_{l_n,n} = o(m_np_n),$$

then

$$\max\{|\theta_{s,n} - \theta_{m_n,n}q_{s,n}| : s = m_n + l_n, \ldots, r_n - m_n\} \to 0.$$

Proof. By Theorem 5.3, we have

$$\begin{aligned}&\max\{|\theta_{s,n} - \theta_{m_n,n}q_{s,n}| : s = m_n + l_n, \ldots, r_n - m_n\} \\ &\leq \; 3(m_np_n)^{-1}\bar{\alpha}_{l_n,n} + 2p_{m_n,n} + (1+(m_np_n)^{-1})p_{l_n,n}.\end{aligned}$$

Since $p_{m_n,n} \leq m_np_n$, Lemma 3.6 implies $p_{l_n,n} = o(m_np_n)$. The Theorem follows. □

Theorem 5.10 *Suppose that* $L = \liminf_{n\to\infty} r_n p_n > 0$*, and let* l_n *and* m_n *be positive integers such that*

$$l_n = o(m_n), \quad m_n p_n \to 0, \quad \bar{\alpha}_{l_n,n} = o(m_n p_n).$$

For every $0 < x < L$*, we have*

$$\theta_{\lceil x/p_n \rceil,n} = \theta_{m_n,n} \exp(-x\theta_{m_n,n}) + o(1).$$

Proof. Define $s_n = \lceil x/p_n \rceil$. Since $x > 0$, we have $m_n = o(s_n)$, and since $x < L$, we have $s_n \le r_n - m_n$. Theorem 5.9 implies that $\theta_{s_n,n} = \theta_{m_n,n} q_{s_n,n} + o(1)$. By Theorem 5.7, we have $q_{s_n,n} = \exp(-s_n \theta_{m_n,n} p_n) + o(1)$. Since $s_n p_n \to x$, the proof is complete. □

Remark 5.11 For $i = 1, \ldots, r_n$, let $I_{i,n}$ be the indicator of the event $A_{i,n}$. Define the random variable

$$T_n = \min\{i \ge 1 : I_{i+1,n} = 1\} \quad \text{conditionally on } I_{1,n} = 1.$$

That is, conditionally on the occurrence of an extreme event on time $i = 1$, the random variable T_n is the waiting time until the next extreme event. The distribution of T_n is

$$\Pr(T_n \ge s \mid A_1) = \Pr\left(\bigcap_{i=1}^{s-1} A_{i+1}^c \,\middle|\, A_1 \right) = \theta_{s,n}.$$

Under the conditions of Theorem 5.10, we have

$$\Pr(p_n T_n \ge x) = \theta_{m_n,n} \exp(-x\theta_{m_n,n}) + o(1), \quad \text{for } x > 0.$$

Hence the normalized inter-arrival time between extreme events ($p_n T_n$) is approximately distributed according to the mixture distribution

$$(1-\theta)\epsilon_0 + \theta \mathrm{Exp}(\theta),$$

where $\theta = \theta_{m_n,n}$, ϵ_0 is the point mass at zero, and $\mathrm{Exp}(\theta)$ is the exponential distribution with mean $1/\theta$.

5.3 Characterization Theorem

The different roles of the $\theta_{m,n}$ in the previous Subsection can be united into a single Characterization Theorem. For sequences a_n and b_n of positive numbers, we write $a_n \asymp b_n$ if

$$0 < \liminf_{n\to\infty} a_n/b_n \le \limsup_{n\to\infty} a_n/b_n < \infty.$$

Theorem 5.12 (Characterization) *Assume that* $r_n p_n \asymp 1$ *and that* $\alpha_{l_n,n} \to 0$ *for some positive integer* l_n *with* $l_n = o(r_n)$*. Let* θ_n *be a sequence of non-negative numbers. The statements (a)–(f) are equivalent:*

(a) There exist positive integers s_n with $s_n \le r_n$ and $s_n \asymp r_n$ such that $q_{s_n,n} = \exp(-s_n\theta_n p_n) + o(1)$.

(b) For every sequence s_n of positive integers with $s_n \le r_n$ and $s_n \asymp r_n$, we have $q_{s_n,n} = \exp(-s_n\theta_n p_n) + o(1)$.

(c) There exist positive integers m_n with $l_n = o(m_n)$, $m_n = o(r_n)$, and $\alpha_{l_n,n} = o(m_n/r_n)$ such that $\theta_{m_n,n} = \theta_n + o(1)$.

(d) For every sequence m_n of positive integers with $l_n = o(m_n)$, $m_n = o(r_n)$, and $\alpha_{l_n,n} = o(m_n/r_n)$, we have $\theta_{m_n,n} = \theta_n + o(1)$.

(e) Same as (c), but with $\theta_{m_n,n}$ replaced by $(m_n p_n)^{-1} p_{m_n,n}$.

(f) Same as (d), but with $\theta_{m_n,n}$ replaced by $(m_n p_n)^{-1} p_{m_n,n}$.

Denote $L = \liminf_{n\to\infty} r_n p_n > 0$. If, additionally, $\limsup_{n\to\infty} \theta_n \le 1$ and $\bar{\alpha}_{l_n,n} \to 0$, then the statements (a)–(f) are also equivalent to each of (g)–(i):

(g) There exists $0 < x \le 1$ with $x < L$ such that $\theta_{\lceil x/p_n \rceil,n} = \theta_n \exp(-x\theta_n) + o(1)$.

(h) There exist $0 < x_1 < x_2 < L$ such that $\theta_{\lceil x_i/p_n \rceil,n} = \theta_n \exp(-x_i\theta_n) + o(1)$ for $i = 1, 2$.

(i) For every $0 < x < L$, we have $\theta_{\lceil x/p_n \rceil,n} = \theta_n \exp(-x\theta_n) + o(1)$.

Proof. *(a) implies (d).* Take positive integers m_n such that $l_n = o(m_n)$, $m_n = o(r_n)$ and $\alpha_{l_n,n} = o(m_n/r_n)$. Since $m_n/s_n \asymp m_n/r_n \to 0$ and $p_{m_n,n} \le m_n p_n \asymp m_n/r_n \to 0$, we obtain by Theorem 5.7 that $q_{s_n,n} = \exp(-s_n\theta_{m_n,n}p_n) + o(1)$, and thus $\exp(-s_n\theta_{m_n,n}p_n) = \exp(-s_n\theta_n p_n) + o(1)$. Since $s_n p_n \asymp r_n p_n \asymp 1$ and $\theta_{m_n,n} \in [0,1]$, we can take logarithms and divide by $s_n p_n$, finding $\theta_{m_n,n} = \theta_n + o(1)$.

(d) implies (c). Since $l_n = o(n)$ and $\alpha_{l_n,n} \to 0$, we can construct a sequence m_n of positive integers such that $l_n = o(m_n)$, $m_n = o(r_n)$, and $\alpha_{l_n,n} = o(m_n/r_n)$; choose for instance m_n such that $m_n/r_n \sim \{\max(l_n/r_n, \alpha_{l_n,n})\}^{1/2}$. By (d), we must also have $\theta_{m_n,n} = \theta_n + o(1)$.

(c) implies (b). Take positive integers s_n such that $s_n \le r_n$ and $s_n \asymp r_n$. We can apply Theorem 5.7 to find

$$q_{s_n,n} = \exp(-s_n\theta_{m_n,n}p_n) + o(1) = \exp\{-s_n[\theta_n + o(1)]p_n\} + o(1) = \exp(-s_n\theta_n p_n) + o(1),$$

where we used that $s_n p_n \asymp 1$.

(b) implies (a). Trivial.

(c) is equivalent to (e), and (d) is equivalent to (f). Since $p_{m_n,n} \le m_n p_n \asymp m_n/r_n \to 0$, we can apply Theorem 5.4(i), obtaining $\theta_{m_n,n} = (m_n p_n)^{-1} p_{m_n,n} + o(1)$.

(d) implies (i). Take $0 < x < L$. Since $l_n = o(r_n)$ and $\bar{\alpha}_{l_n,n} \to 0$, we can find a sequence m_n of positive integers such that $l_n = o(m_n)$, $m_n = o(r_n)$, and $\bar{\alpha}_{l_n,n} = (m_n/r_n)$. By Theorem 5.10 and by (d), we have

$$\begin{aligned}\theta_{\lceil x/p_n \rceil,n} &= \theta_{m_n,n}\exp(-x\theta_{m_n,n}) + o(1) \\ &= [\theta_n + o(1)]\exp\{-x[\theta_n + o(1)]\} + o(1) = \theta_n \exp(-x\theta_n) + o(1).\end{aligned}$$

(i) implies (g) and (h). Trivial.

(g) implies (c). As before, we can find integers $m_n \geq 1$ such that $l_n = o(m_n)$, $m_n = o(r_n)$, and $\alpha_{l_n,n} \leq \bar{\alpha}_{l_n,n} = o(m_n/r_n)$. By Theorem 5.10, we have

$$\theta_{m_n,n} \exp(-x\theta_{m_n,n}) = \theta_n \exp(-x\theta_n) + o(1).$$

Without loss of generality, we can restrict attention to subsequences along which $\theta_n \to \theta \in [0,1]$ and $\theta_{m_n,n} \to \theta' \in [0,1]$. Clearly $\theta \exp(-x\theta) = \theta' \exp(-x\theta')$. Since the function $z \mapsto z \exp(-xz)$ is strictly increasing in $z \in [0, 1/x]$, and since $1/x \geq 1$, we have $\theta = \theta'$.

(h) implies (c). There exist integers $m_n \geq 1$ with $l_n = o(m_n)$, $m_n = o(r_n)$, and $\alpha_{l_n,n} \leq \bar{\alpha}_{l_n,n} = o(m_n/r_n)$ such that

$$\theta_{m_n,n} \exp(-x_i\theta_{m_n,n}) = \theta_n \exp(-x_i\theta_n) + o(1) \quad \text{for } i = 1, 2.$$

If $\theta_n \to \theta \in [0,1]$ and $\theta_{m_n,n} \to \theta' \in [0,1]$ along some subsequence, then

$$\theta \exp(-x_i\theta) = \theta' \exp(-x_i\theta'), \quad \text{for } i = 1, 2.$$

If $\theta = 0$, then $\theta' \exp(-x_i\theta') = 0$, and thus $\theta' = 0$ [in fact, here we only need one single $0 < x < L$]. If $\theta > 0$, then either (1) $\theta = \theta'$ or (2) $\theta \neq \theta'$ and $(\theta - \theta')^{-1} \log(\theta/\theta') = x_i$ for $i = 1, 2$. Since $x_1 < x_2$, the second case is impossible, and thus $\theta = \theta'$. □

6 Application: intervals estimator

Two popular estimators for the extremal index are the blocks and the runs estimator (Hsing 1991 and 1993). Both of them require the choice of a tuning parameter, which, unfortunately, often has a grave impact on the final estimates. Ferro and Segers (2002) used the asymptotic distribution of the random times between threshold exceedances to construct the so-called intervals estimator, for which no such choice must be made. Consistency of the estimator was demonstrated under the stringent condition that the sequence of random variables is m-dependent. This assumption, however, is unnecessarily restrictive, as will be shown next in our general setting.

In the finite-sample case, let $A_1, \ldots, A_n$ be block-stationary events. Denote $I_i = I(A_i)$, the indicator of A_i, and let $N = \sum_{i=1}^n I_i$ be the number of events occurred. Put $S_0 = 0$, $S_{N+1} = n + 1$, and in case $N \geq 1$ let $1 \leq S_1 < \cdots < S_N \leq n$ be the times at which events occurred, that is, $\{i = 1, \ldots, n : I_i = 1\} = \{S_1, \ldots, S_N\}$. Denote the inter-arrival times by $T_i = S_{i+1} - S_i$, for $i = 0, \ldots, N$.

The intervals estimator is based on the statistic

$$\tau = \sum_{i=1}^n \sum_{j=i}^n \prod_{k=i}^j (1 - I_k) = \sum_{t=0}^N \frac{1}{2}(T_t - 1)T_t$$

with expectation

$$E(\tau) = \sum_{i=1}^n \sum_{j=i}^n \Pr\left(\bigcap_{k=i}^j A_k^c\right) = \sum_{s=1}^n (n - s + 1)q_s.$$

If we, naively, plug in the approximation $q_s \approx \exp(-sp\theta_m)$, see Theorem 5.7, then we may guess that $E(\tau) \approx n/(p\theta_m)$. Hence, given an estimator $\hat{p}$ of p, we may estimate θ_m by $\hat{\theta}_m = n/(\hat{p}\tau)$, a variant of the intervals estimator of Ferro and Segers (2002). A possible candidate for $\hat{p}$ is of course N/n.

The asymptotic theory to follow requires an upper bound for $\mathrm{Var}\,(\tau)$. As usual, denote $\bar{\alpha}_l = \max\{\alpha_{s,l} : s = l, \ldots, n\}$, with $\alpha_{s,l}$ the mixing coefficients of Section 2.

Lemma 6.1 *For integer $1 \leq l \leq n$, we have*

$$\mathrm{Var}\,(\tau) \leq 2n \sum_{s=1}^{n} (s + 2l) s q_s + n^4 \bar{\alpha}_l.$$

Proof. Denoting

$$\begin{aligned} A &= \{(i,j,u,v) \in \{1,\ldots,n\}^4 : i \leq j,\, u \leq v\} \\ C(i,j,u,v) &= \mathrm{Cov}\left(\prod_{k=i}^{j} (1 - I_k), \prod_{w=u}^{v} (1 - I_w) \right), \quad \text{for } (i,j,u,v) \in A, \end{aligned}$$

we have $\mathrm{Var}\,(\tau) = \sum_A C(i,j,u,v)$. Now for $\nu = 0, 1, \ldots, 6$, let A_ν be the set of all $(i,j,u,v) \in A$ such that

case $\nu = 0$	: $i = u$;		
case $\nu = 1$	: $i < u \leq j$;	case $\nu = 4$	: $u < i \leq v$;
case $\nu = 2$	: $j < u \leq j + l$;	case $\nu = 5$	: $v < i \leq v + l$;
case $\nu = 3$	: $j + l < u$;	case $\nu = 6$	: $v + l < i$.

The sets $A_0, \ldots, A_6$ form a partition of A, hence

$$\mathrm{Var}\,(\tau) = \sum_{\nu=0}^{6} \sum_{A_\nu} C(i,j,u,v) = \sum_{A_0} C(i,j,u,v) + 2 \sum_{\nu=1}^{3} \sum_{A_\nu} C(i,j,u,v),$$

by symmetry. On A_0, we have $C(i,j,u,v) \leq q_{\max(j,v)-i+1}$, hence

$$\sum_{A_0} C(i,j,u,v) \leq \sum_{s=1}^{n} q_s \sum_{A_0} \mathbf{1}_{\{\max(j,v)-i+1=s\}} \leq 2n \sum_{s=1}^{n} s q_s.$$

On A_1 as well, we have $C(i,j,u,v) \leq q_{\max(j,v)-i+1}$, hence

$$\sum_{A_1} C(i,j,u,v) \leq \sum_{s=1}^{n} q_s \sum_{A_1} \mathbf{1}_{\{\max(j,v)-i+1=s\}} \leq n \sum_{s=1}^{n} (s-1) s q_s.$$

On A_2, we have $C(i,j,u,v) \leq q_{\max(j-i+1,v-u+1)}$, so that

$$\sum_{A_2} C(i,j,u,v) \leq \sum_{s=1}^{n} q_s \sum_{A_2} \mathbf{1}_{\{\max(j-i+1,v-u+1)=s\}} \leq 2nl \sum_{s=1}^{n} s q_s.$$

Finally, on A_3, we have $C(i,j,u,v) \le \bar{\alpha}_l$, and thus

$$\sum_{A_3} C(i,j,u,v) \le \frac{1}{2} n^4 \bar{\alpha}_l.$$

To conclude the proof, add the bounds on $\sum_{A_\nu} C(i,j,u,v)$. $\square$

Next we consider the asymptotic case. For $n \ge 1$, let $A_{1,n}, \ldots, A_{n,n}$ be block-stationary events, and for $i = 1, \ldots, n$, let the random variable $I_{i,n}$ be the indicator of the event $A_{i,n}$. The statistic of interest is

$$\tau_n = \sum_{i=1}^{n} \sum_{j=i}^{n} \prod_{k=i}^{j} (1 - I_{k,n}).$$

Theorem 6.2 *Let $1 \le l_n \le m_n \le n$ be integers. If*

$$l_n = o(m_n), \quad p_{m_n,n} \to 0, \quad m_n = o(np_{m_n,n}), \quad \textit{and} \quad \alpha_{l_n,n} = o(m_n/n).$$

then

$$E(\tau_n) \sim \frac{m_n n}{p_{m_n,n}} \sim \frac{n}{p_n \theta_{m_n,n}}.$$

Proof. Since $\alpha_{l_n,n} = o(p_{m_n,n})$, we have, according to Theorem 5.4, $\theta_{m_n,n} \sim (m_n p_n)^{-1} p_{m_n,n}$, which proves the second asymptotic equivalence.

We prove the first asymptotic equivalence by separately considering the lim sup and the lim inf. By the upper bound in Lemma 3.2, we have

$$E(\tau_n) \le m_n n + n \sum_{s=m_n+1}^{n} q_{s,n} \le m_n n + n \sum_{s=m_n+1}^{n} q_{m_n,n}^{k} + n^2 \alpha_{l_n,n}/p_{m_n,n},$$

where $k = \lfloor (s + l_n)/(m_n + l_n) \rfloor$. The conditions imply

$$E(\tau_n) \le n \sum_{s=m_n+1}^{n} q_{m_n,n}^{k} + o(m_n n/p_{m_n,n}).$$

Since $k = \lfloor (s + l_n)/(m_n + l_n) \rfloor > (s - m_n)/(m_n + l_n)$, we have

$$\frac{k}{s/m_n} > \frac{s - m_n}{s} \frac{m_n}{m_n + l_n} > 1 - \frac{m_n}{s} - \frac{l_n}{m_n}.$$

Take $0 < \epsilon < 1$. For n large enough so that $m_n/n < \epsilon/2$ and $l_n/m_n < \epsilon/2$, we have $k/(s/m_n) > 1 - \epsilon$ for $s = \lceil 2m_n/\epsilon \rceil, \ldots, n$. Hence

$$\begin{aligned}
E(\tau_n) &\le 2m_n n/\epsilon + n \sum_{s=\lceil 2m_n/\epsilon \rceil}^{n} q_{m_n,n}^{(1-\epsilon)s/m_n} + o(m_n n/p_{m_n,n}) \\
&\le n \sum_{s=0}^{\infty} q_{m_n,n}^{(1-\epsilon)s/m_n} + o(m_n n/p_{m_n,n}) = \frac{n}{1 - q_{m_n,n}^{(1-\epsilon)/m_n}} + o(m_n n/p_{m_n,n}) \\
&\le \frac{m_n n}{(1-\epsilon) p_{m_n,n}} + o(m_n n/p_{m_n,n}).
\end{aligned}$$

Let $\epsilon \downarrow 0$ to find $\limsup_{n\to\infty} p_{m_n,n}E(\tau_n)/(m_n n) \le 1$.

Next, we deal with the lim inf. Let $0 < \epsilon < 1$. For large enough n, we have $m_n/(\epsilon p_{m_n,n}) < n$. Set $a_n = \lceil m_n/\epsilon \rceil$ and $b_n = \lceil m_n/(\epsilon p_{m_n,n}) \rceil$. We have

$$E(\tau_n) \ge \sum_{s=a_n}^{b_n} (n-s+1) q_{s,n} \ge (n-b_n+1) \sum_{s=a_n}^{b_n} q_{s,n}.$$

Set $k = \lceil (s+l_n)/(m_n+l_n) \rceil$ for $s = a_n, \ldots, b_n$. By Lemma 3.2, we have

$$\begin{aligned} E(\tau_n) &\ge (n-b_n+1) \sum_{s=a_n}^{b_n} \{q_{m_n,n}^k - (\alpha_{l_n,n} + p_{l_n,n})/p_{m_n,n}\}. \\ &\ge (n-b_n+1) \sum_{s=a_n}^{b_n} q_{m_n,n}^k - \frac{nm_n}{\epsilon p_{m_n,n}^2}(\alpha_{l_n,n} + p_{l_n,n}). \end{aligned}$$

Since $\alpha_{m_n,n} = o(m_n/n)$ and $m_n/n = o(p_{m_n,n})$, we have $\alpha_{l_n,n} = o(p_{m_n,n})$, so that also $p_{l_n,n} = o(p_{m_n,n})$ according to Lemma 3.6(ii). Hence

$$E(\tau_n) \ge (n-b_n+1) \sum_{s=a_n}^{b_n} q_{m_n,n}^k + o(m_n n/p_{m_n,n}).$$

Now $k = \lceil (s+l_n)/(m_n+l_n) \rceil \le (s+2l_n+m_n)/(m_n+l_n) \le (s+l_n+m_n)/m_n$, so that $k/(s/m_n) \le 1 + 2m_n/s \le 1 + 2\epsilon$ for $s \ge a_n$. Consequently,

$$\begin{aligned} E(\tau_n) &\ge (n-b_n+1) \sum_{s=a_n}^{b_n} q_{m_n,n}^{(1+2\epsilon)s/m_n} + o(m_n n/p_{m_n,n}) \\ &= (n-b_n+1) \frac{q_{m_n,n}^{(1+2\epsilon)a_n/m_n} - q_{m_n,n}^{(1+2\epsilon)(b_n+1)/m_n}}{1 - q_{m_n,n}^{(1+2\epsilon)/m_n}} + o(m_n n/p_{m_n,n}). \end{aligned}$$

Now we have

$$\begin{aligned} q_{m_n,n}^{(1+2\epsilon)a_n/m_n} &\to 1, \\ q_{m_n,n}^{(1+2\epsilon)(b_n+1)/m_n} &\to \exp\{-(1+2\epsilon)/\epsilon\}, \\ 1 - q_{m_n,n}^{(1+2\epsilon)/m_n} &\sim (1+2\epsilon)p_{m_n,n}/m_n. \end{aligned}$$

Hence $\liminf_{n\to\infty} p_{m_n,n}E(\tau_n)/(m_n n) \ge [1 - \exp\{-(1+2\epsilon)/\epsilon\}]/(1+2\epsilon)$. Let $\epsilon \downarrow 0$ to conclude the proof. □

Remark 6.3 The inequalities $m_n p_n \theta_{m_n,n} \le p_{m_n,n} \le m_n p_n$ (see Theorem 5.1) yield simple sufficient conditions for Theorem 6.2: first, $m_n p_n \to 0$ implies $p_{m_n,n} \to 0$; second, in the typical case $\liminf_{n\to\infty} \theta_{m_n,n} > 0$, the condition $np_n \to \infty$ implies $m_n = o(np_{m_n,n})$.

Theorem 6.4 *If $\bar{\alpha}_{l_n,n} = o(m_n^2/(n^2 p_{m_n,n}^2))$ in addition to the conditions of Theorem 6.2, then* $\mathrm{Var}\,(\tau_n) = o(n^2 m_n^2/p_{m_n,n}^2)$, *and hence*

$$\frac{p_n \theta_{m_n,n}}{n} \tau_n \to 1 \quad \textit{in } L^2.$$

In particular, if $\hat{p}_n = p_n\{1 + o_p(1)\}$, *then*

$$\hat{\theta}_n = n/(\hat{p}_n \tau_n) = \theta_{m_n,n}\{1 + o_p(1)\}.$$

Proof. By assumption, we have $n^4 \bar{\alpha}_{l_n,n} = o(n^2 m_n^2/p_{m_n,n}^2)$, so that by Lemma 6.1 it is sufficient to show that $\sum_{s=1}^n (s + 2l_n) s q_{s,n} = o(n m_n^2/p_{m_n,n}^2)$. Now, for n large enough so that $2l_n \le m_n$, we have

$$\sum_{s=1}^{n} (s + 2l_n) s q_{s,n} \le 2m_n^2 + 2 \sum_{s=m_n}^{n} s^2 q_{s,n}.$$

Clearly, we may restrict attention to the second term on the right-hand side of this inequality. Set $a_n = \lceil 2m_n/p_{m_n} \rceil$. Since $m_n = o(np_{m_n,n})$, we have $a_n \le n$ for large enough n. So we can write

$$\sum_{s=m_n}^{n} s^2 q_{s,n} \le \sum_{s=m_n}^{a_n - 1} s^2 q_{s,n} + \sum_{s=a_n}^{n} s^2 q_{s,n} = I_n + II_n,$$

say. By Lemma 3.2, we have

$$I_n \le \sum_{s=m_n}^{a_n-1} s^2 \left(q_m^{\lfloor (s+l)/(m+l) \rfloor} + \frac{\alpha_{l_n,n}}{p_{m_n,n}} \right) \le \sum_{s=m_n}^{a_n-1} s^2 q_m^{s/(4m_n)} + a_n^3 \frac{\alpha_{l_n,n}}{p_{m_n,n}}.$$

Since $\sum_{s=1}^{\infty} s^2 (1-\epsilon)^s = O(\epsilon^{-3})$ as $0 < \epsilon \to 0$, and since $\alpha_{l_n,n} = o(m_n/n)$, we have

$$\begin{aligned} I_n &= O\left(\left(1 - q_{m_n,n}^{1/(4m_n)}\right)^{-3} \right) + O\left(m_n^3 \alpha_{l_n,n}/p_{m_n,n}^4 \right) \\ &= O(m_n^3/p_{m_n,n}^3) + o\left(m_n^4/(np_{m_n,n}^4) \right). \end{aligned}$$

Moreover, $m_n = o(np_{m_n,n})$, so that $I_n = o(nm_n^2/p_{m_n,n}^2)$.

Next, we deal with II_n. By Lemma 3.2, we have

$$II_n \le \sum_{s=a_n}^{n} s^2 \left(q_{a_n,n}^{\lfloor (s+l_n)/(a_n+l_n) \rfloor} + \frac{\alpha_{l_n,n}}{p_{a_n,n}} \right) \le \sum_{s=a_n}^{n} s^2 q_{a_n,n}^{s/(4a_n)} + n^3 \frac{\alpha_{l_n,n}}{p_{a_n,n}}.$$

Apply Lemma 3.2 again to find

$$q_{a_n,n} \le q_{m_n,n}^{\lfloor (a_n+l_n)/(m_n+l_n) \rfloor} + \frac{\alpha_{l_n,n}}{p_{m_n,n}} \le q_{m_n,n}^{a_n/(4m_n)} + o(1) \to \exp(-1/2).$$

Hence we can find a number $0 < \delta < 1$ such that $q_{a_n,n} \le 1 - \delta$ for all large enough n. We obtain

$$II_n = O\left([1 - (1-\delta)^{1/(4a_n)}]^{-3} \right) + O(n^3 \alpha_{l_n,n}).$$

Since $[1 - (1-\delta)^{1/4a_n}]^{-3} \sim (4/\delta)^3 a_n^3 = O(m_n^3/p_n^3)$ and $\alpha_{l_n,n} = o(m_n^2/(n^2 p_{m_n,n}^2))$, we conclude $II_n = o(nm_n^2/p_n^2)$. □

7 Multiple extreme events

In a multivariate time series there are different forms of dependence to consider, such as the dependence between the marginals at a fixed time point and the dependence over time in each of the marginal series. However, the exceptional events in each of the marginals may also depend on one another in a more complicated way.

Example 7.1 Let $\{Y_n : n \geq 1\}$ be independent and identically distributed random variables, and consider the stationary bivariate time series $\mathbf{X}_n = (X_n^{(1)}, X_n^{(2)}) = (Y_n, Y_{n+1})$, for $n \geq 1$. For each n the marginal variables $X_n^{(1)}$ and $X_n^{(2)}$ are independent, and each of the marginal time series $\{X_n^{(i)} : n \geq 1\}$ consists of independent random variables. Nevertheless, the coordinate-wise maxima $M_n^{(i)} = \max_{j=1,\dots,n} X_j^{(i)}$ satisfy

$$\Pr\left(M_n^{(1)} \leq u_n^{(1)},\, M_n^{(2)} \leq u_n^{(2)}\right) = \Pr\left(M_n^{(1)} \leq \min(u_n^{(1)}, u_n^{(2)})\right) + o(1)$$

for any sequence $\{(u_n^{(1)}, u_n^{(2)})\}$, that is, $M_n^{(1)}$ and $M_n^{(2)}$ are completely dependent in the limit.

For every $n \geq 1$, let $A_{1,n}, \dots, A_{r_n,n}$ and $B_{1,n}, \dots, B_{r_n,n}$ be events on a common probability space (which may vary with n). Define $C_{i,n} = A_{i,n} \cup B_{i,n}$ for $n \geq 1$ and $i = 1, \dots, r_n$. For $Z = A, B, C$, assume that the events $Z_{1,n}, \dots, Z_{r_n,n}$ are block-stationary, and put

$$p_{m,n}^Z = \Pr\left(\bigcup_{i=1}^m Z_{i,n}\right), \qquad p_n^Z = p_{1,n}^Z,$$

$$q_{m,n}^Z = 1 - p_{m,n}^Z, \qquad \theta_{m,n}^Z = \Pr\left(\bigcap_{i=2}^m Z_{i,n}^c \,\middle|\, Z_{1,n}\right),$$

where $m = 1, \dots, r_n$. Define the mixing coefficients

$$\alpha_{s,l,n} = \max_{Z=A,B,C} \max\Big\{\left|\Pr\left(\bigcap_{i=u+1}^{v} Z_{i,n}^c \cap \bigcap_{j=s+v+1}^{s+w} Z_{j,n}^c\right) - q_{v-u,n}^Z q_{w-v,n}^Z\right| : u \geq 0, v-u \geq l, w-v \geq l, w+s \leq r_n\Big\},$$

with $\alpha_{s,l,n} = 0$ if $2l + s > r_n$. Abbreviate $\alpha_{l,n} = \alpha_{l,l,n}$.

We will investigate the dependence between the A-array and the B-array through the quantity

$$\theta_{m,n}^{A|B} = \Pr\left(\bigcap_{i=1}^m A_{i+j,n}^c \,\middle|\, \bigcup_{i=1}^m B_{i+j,n}\right) = (p_{m,n}^C - p_{m,n}^A)/p_{m,n}^B$$

where $m = 1, \dots, r_n$ and $j = 0, \dots, r_n - m_n$. Although $\theta_{m,n}^{A|B}$ and $\theta_{m,n}^{B|A}$ are not the same, any statement on $\theta_{m,n}^{A|B}$ obviously corresponds to another one with the roles of A and B interchanged.

Theorem 7.2 *For positive integers l_n and m_n such that*

$$l_n = o(m_n), \quad m_n = o(r_n), \quad \textit{and} \quad \alpha_{l_n,n} = o(\max(m_n/r_n, p^C_{m_n,n})),$$

we have

$$q^C_{r_n,n} = q^A_{r_n,n} \left(q^B_{r_n,n}\right)^{\theta^{A|B}_{m_n,n}} + o(1).$$

Proof. By Theorem 3.7 and since $r_n/m_n \to \infty$, we have

$$\begin{aligned} q^C_{r_n,n} &= \left(q^C_{m_n,n}\right)^{r_n/m_n} + o(1) = \exp\{-(r_n/m_n)p^C_{m_n,n}\} + o(1) \\ &= \exp\{-(r_n/m_n)(p^A_{m_n,n} + p^B_{m_n,n}\theta^{A|B}_{m_n,n})\} + o(1). \end{aligned}$$

Without loss of generality, we may assume that $(r_n/m_n)p^C_{m_n,n} \to \lambda \in [0,\infty]$.

Assume first that $\lambda < \infty$. Then $\alpha_{l_n,n} = o(m_n/r_n)$, so that by Theorem 3.7

$$q^Z_{r_n,n} = \exp\{-(r_n/m_n)p^Z_{m_n,n}\} + o(1), \quad \text{for } Z = A, B.$$

Since $\liminf q^B_{r_n,n} \geq \liminf q^C_{r_n,n} = \exp(-\lambda) > 0$, we obtain

$$q^C_{r_n,n} = \{q^A_{r_n,n} + o(1)\}\{q^B_{r_n,n} + o(1)\}^{\theta^{A|B}_{m_n,n}} + o(1) = q^A_{r_n,n} \left(q^B_{r_n,n}\right)^{\theta^{A|B}_{m_n,n}} + o(1).$$

Next, assume that $\lambda = \infty$. Then $q^C_{r_n,n} \to 0$. Since $p^C_{m,n} = p^A_{m,n} + p^B_{m,n}\theta^{A|B}_{m,n}$, we can without loss of generality restrict n to a further subsequence for which

$$\liminf p^A_{m_n,n}/p^C_{m_n,n} > 0 \quad \text{or} \quad \liminf p^B_{m_n,n}\theta^{A|B}_{m_n,n}/p^C_{m_n,n} > 0.$$

In the first case we have $\alpha_{l_n,n} = o(p^A_{m_n,n})$ and thus, by Theorem 3.7,

$$q^A_{r_n,n} = \exp\{-(r_n/m_n)p^A_{m_n,n}\} + o(1) \to 0.$$

In the second case we have $\alpha_{l_n,n} = o(p^B_{m_n,n})$ and thus, by Theorem 3.7,

$$q^B_{r_n,n} = \exp\{-(r_n/m_n)p^B_{m_n,n}\} + o(1) \to 0.$$

As $p^B_{m,n} \leq p^C_{m,n}$, we have $\liminf \theta^{A|B}_{m_n,n} > 0$, and thus $\left(q^B_{r_n,n}\right)^{\theta^{A|B}_{m_n,n}} \to 0$. □

The dependence coefficient $\theta^{A|B}_{m,n}$ is related to the indices $\theta^Z_{m,n}$ for $Z = A, B, C$.

Theorem 7.3 *Let l_n and m_n be positive integers with $2m_n + l_n \leq r_n$. If*

$$l_n = o(m_n), \quad p^C_{m_n,n} \to 0, \quad \textit{and} \quad \alpha_{l_n,n} = o(m_n p^C_n),$$

then

$$p^A_n\theta^A_{m_n,n} + p^B_n\theta^B_{m_n,n}\theta^{A|B}_{m_n,n} = p^C_n\left[\theta^C_{m_n,n} + o(1)\right].$$

Proof. By Theorem 5.4, we have

$$\theta^C_{m_n,n} = \frac{p^C_{m_n,n}}{m_n p^C_n} + o(1) = \frac{p^A_{m_n,n}}{m_n p^C_n} + \frac{p^B_{m_n,n}}{m_n p^C_n}\theta^{A|B}_{m_n,n} + o(1).$$

Without loss of generality, we can restrict n to a subsequence along which $p^A_n/p^C_n \to \lambda \in [0,1]$ and $p^B_n/p^C_n \to \mu \in [0,1]$. If $\lambda = 0$, then

$$\frac{p^A_{m_n,n}}{m_n p^C_n} \leq \frac{p^A_n}{p^C_n} \to 0,$$

while if $\lambda > 0$, then $\alpha_{l_n,n} = o(m_n p^A_n)$ and thus, by Theorem 5.4,

$$p^A_{m_n,n} = m_n p^A_n \{\theta^A_{m_n,n} + o(1)\}.$$

The arguments for the B-term are analogous. □

The value of $\theta^{A|B}_{m,n}$ is approximately the same for a range of values of m.

Theorem 7.4 *Let l_n, m_n, and M_n be positive integers such that $l_n \leq m_n \leq M_n$ and $2M_n + l_n \leq r_n$. If*

$$l_n = o(m_n), \quad \alpha_{l_n,n} = o(m_n p^C_n), \quad p^C_{M_n,n} \to 0, \quad M_n p^C_n = O(1),$$

and $\liminf p^B_n \theta^B_{M_n,n}/p^C_n > 0$, *then*

$$\begin{aligned} \theta^{A|B}_{m_n,n} &= \left(p^C_n \theta^C_{m_n,n} - p^A_n \theta^A_{m_n,n}\right)/p^B_n \theta^B_{m_n,n} + o(1) \\ &= \left(p^C_n \theta^C_{M_n,n} - p^A_n \theta^A_{M_n,n}\right)/p^B_n \theta^B_{M_n,n} + o(1) = \theta^{A|B}_{M_n,n} + o(1). \end{aligned}$$

Proof. By Theorem 7.3, we have

$$\begin{aligned} \theta^{A|B}_{m_n,n} &= \frac{p^C_n \left[\theta^C_{m_n,n} + o(1)\right] - p^A_n \theta^A_{m_n,n}}{p^B_n \theta^B_{m_n,n}}, \\ \theta^{A|B}_{M_n,n} &= \frac{p^C_n \left[\theta^C_{M_n,n} + o(1)\right] - p^A_n \theta^A_{M_n,n}}{p^B_n \theta^B_{M_n,n}}. \end{aligned}$$

Since $\liminf p^B_n \theta^B_{m_n,n}/p^C_n > 0$, we have $\liminf p^B_n/p^C_n > 0$ and $\liminf \theta^B_{m_n,n} > 0$. Therefore $\alpha_{l_n,n} = o(m_n p^B_n)$ and thus, by Theorem 5.6,

$$\theta^B_{m_n,n} = \theta^B_{M_n,n} + o(1) \quad \text{and} \quad \theta^C_{m_n,n} = \theta^C_{M_n,n} + o(1).$$

Hence, for subsequences along which $p^A_n/p^C_n \to 0$, we have

$$\theta^{A|B}_{m_n,n} = \frac{p^C_n \theta^C_{m_n,n}}{p^B_n \theta^B_{m_n,n}} + o(1) = \frac{p^C_n \theta^C_{M_n,n}}{p^B_n \theta^B_{M_n,n}} + o(1) = \theta^{A|B}_{M_n,n}.$$

On the other hand, for subsequences along which $\liminf p^A_n/p^C_n > 0$, we have, by Theorem 5.6, $\theta^A_{m_n,n} = \theta^A_{M_n,n} + o(1)$, leading to the stated expression. □

Remark 7.5 The previous results suggest three ways to estimate the dependence coefficient $\theta_{m,n}^{A|B}$ when observing the indicator variables $I(A_{i,n})$ and $I(B_{i,n})$ for $i = 1, \ldots, n$: (1) estimate $p_{m,n}^Z$ for $Z = A, B, C$ and use the definition; (2) estimate $q_{r,n}^Z$ for $Z = A, B, C$ and use Theorem 7.2; (3) estimate $p_n^Z \theta_{m,n}^Z$ for $Z = A, B, C$ and use Theorem 7.4. In (1) and (2) one could employ the disjoint-blocks estimator of Section 4, while in (3) the intervals estimator of Section 6 for the extremal indices $\theta_{m,n}^Z$ would lead to an estimator of $\theta_{m,n}^{A|B}$ for which no block length needs to be chosen. The properties of these estimators remain open for further research.

References

Ferro, C.A.T. and Segers, J. (2002) Inference for clusters of extremes. *J. Roy. Statist. Soc. Ser. B*, to appear.

Gumbel, E.J. (1958) *Statistics of Extremes.* Columbia University Press, New York.

Hsing, T. (1989) Extreme value theory for multivariate stationary sequences. *J. Multivariate Anal.*, **29**, 274–291.

Hsing, T. (1991) Estimating the parameters of rare events. *Stochastic Process. Appl.*, **37**, 117–139.

Hsing, T. (1993) Extremal index estimation for a weakly dependent stationary sequence. *Ann. Statist.*, **21**, 2043–2071.

Hsing, T., Hüsler, J., and Leadbetter, M.R. (1988) On the exceedance point process for a stationary sequence. *Probab. Theory Related Fields*, **78**, 97–112.

Leadbetter, M.R. (1974) On extreme values in stationary sequences. *Z. Wahrsch. Verw. Gebiete*, **28**, 289–303.

Leadbetter, M.R. (1983) Extremes and local dependence of stationary sequences. *Z. Wahrsch. Verw. Gebiete*, **65**, 291–306.

Leadbetter, M.R., Lindgren, G., and Rootzén, H. (1983) *Extremes and Related Properties of Random Sequences and Processes.* Springer-Verlag, New York.

Nandagopalan, S. (1994) On the multivariate extremal index, *J. Res. Natl. Inst. Stand. Technol.*, **99**, 543–550.

O'Brien, G.L. (1987) Extreme values for stationary and Markov sequences. *Ann. Probab.*, **15**, 281–291.

Smith, R.L. and Weissman, I. (1994) Estimating the extremal index. *J. Roy. Statist. Soc. Ser. B*, **56**, 515–528.

Smith, R.L. and Weissman, I. (1996) *Characterization and Estimation of the Multivariate Extremal Index.* Draft, available at www.unc.edu/depts/statistics/postscript/rs/extremal.ps.