Statistics and Probability Letters 137 (2018) 124–134

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

On the maxima and minima of complete and incomplete samples from nonstationary random fields

Zacarias Panga, Luísa Pereira *

Centro de Matemática e Aplicações da Universidade da Beira Interior (CMA-UBI), Portugal

A R T I C L E   I N F O

Article history:
Received 2 December 2017
Received in revised form 17 January 2018
Accepted 19 January 2018
Available online 31 January 2018

Keywords:
Random field
Incomplete sample
Dependence conditions

A B S T R A C T

We derive the joint limiting distribution for the maxima and minima of complete and incomplete samples from nonstationary random fields satisfying appropriate dependence conditions. The results are illustrated with a nonstationary Gaussian random field.

1. Introduction

The extreme value analysis for randomly or deterministic missing values is a very recent research field. The unavailability of data may be caused, for instance, by failures of measuring systems. In the study of the properties of a process with failures it may be crucial to take into account the loss of information so, the effect of deterministic or randomly missing values on extremes of stationary sequences has been extensively studied (see, e.g., Ferreira, 1995, Hall, 2006; Hall and Temido, 2008, Hashorva and Weng, 2014, Mladenović and Piterbarg, 2006, Robert, 2010, Tan and Wang, 2012, Peng et al., 2011, Weissman and Cohen, 1995, and references therein).

So far many results have been obtained for extremes of incomplete sequences, the effect of missing values in extremes of a random field has not been investigated. Random fields are of increasing interest in applications such as Environmental Assessment over entire regions of space and constitute an area of current research. A considerable amount of work has been done in extending results of extreme value theory to random fields whose variables are located on a regular grid identified with $\mathbb{Z}^2$ (see, e.g., Choi, 2010, Ferreira and Pereira, 2008, 2012, Leadbetter and Rootzén, 1998, Pereira and Tan, 2017; Pereira, 2013; Pereira and Ferreira, 2006; Pereira, 2010). Although these important advances in the theory of extremes of random fields over $\mathbb{Z}^2$, it was not obvious how to apply the results to non-gridded observations. Recently, Pereira et al. (2017) extended the existing theory, concerning the asymptotic behavior of the maximum and the extremal index to random fields defined over discrete subsets of $\mathbb{R}^2$.

Following the idea of Pereira et al. (2017) in this paper we extend the results of Peng et al. (2011), concerning the asymptotic behavior of maxima and minima of complete and incomplete samples of stationary sequences of random variables, to nonstationary random fields, $\mathbf{Z} = \{Z(x) : x \in S\}$, where $S = \bigcup_{n\geq 1} A_n$ and $A = \{A_n\}_{n\geq 1}$ is an increasing sequence of sets of isolated points of $\mathbb{R}^2$, subject to dependence conditions. We will assume, without loss of generality, that the variables $Z(x)$, $x \in S$, have common distribution F, being $\bar{F}$ the corresponding survival function. We will denote the maximum and the minimum of $Z(x)$ over $B \subset S$ by $\bigvee_{x\in B} Z(x)$ and $\bigwedge_{x\in B} Z(x)$, respectively.

* Corresponding author.
E-mail address: lpereira@ubi.pt (L. Pereira).

https://doi.org/10.1016/j.spl.2018.01.019

In what concerns to the probabilistic model governing the missing observations we shall consider independent Bernoulli random variables, $\epsilon_n(x)$, $x \in A_n$, $n \geq 1$, independent of $Z(x)$ so that $\epsilon_n(x)$ is the indicator of the event that $Z(x)$, $x \in A_n$, is observed. Thus $S(A_n) = \sum_{x\in A_n} \epsilon_n(x)$ is just the number of observed random variables from $\{Z(x) : x \in A_n\}$. The restriction on $\epsilon_n(x)$, $x \in A_n$ is that, as $n \to \infty$,

$$\frac{S(A_n)}{f(n)} \xrightarrow{P} p \in \,]0, 1\,], \tag{1.1}$$

where $f(n)$ denotes the cardinality of A_n, i.e., $f(n) = \sharp A_n$.

For an incomplete sample of an arbitrary subset $B_n \subset S$, we define the maximum and minimum as follows

$$\widetilde{\bigvee}_{x\in B_n} Z(x) = \begin{cases} \bigvee_{x\in B_n,\ \epsilon(x)=1} Z(x) & , \ S(B_n) \geq 1 \\ \inf\{t : F_Z(t) > 0\} & , \ S(B_n) = 0. \end{cases}$$

$$\widetilde{\bigwedge}_{x\in B_n} Z(x) = \begin{cases} \bigwedge_{x\in B_n,\ \epsilon(x)=1} Z(x) & , \ S(B_n) \geq 1 \\ \inf\{t : F_Z(t) > 0\} & , \ S(B_n) = 0. \end{cases}$$

The paper is organized as follows. We introduce in Section 2 a coordinatewise mixing condition which allows us to obtain the asymptotic independence of jointly maxima and minima over disjoint discrete subsets of $\mathbb{R}^2$. In Section 3 we derive the joint asymptotic distributions of the normalized random vectors $\left(\bigvee_{x\in S} Z(x), \bigwedge_{x\in S} Z(x), \widetilde{\bigvee}_{x\in S} Z(x), \widetilde{\bigwedge}_{x\in S} Z(x)\right)$, $\left(\bigvee_{x\in S} Z(x), \bigwedge_{x\in S} Z(x)\right)$ and $\left(\widetilde{\bigvee}_{x\in S} Z(x), \widetilde{\bigwedge}_{x\in S} Z(x)\right)$ considering additionally to the coordinatewise mixing condition, the local dependence condition of Davis (1979) tailored for random fields, which implies that there cannot be clustering of high and low values of the random field $\mathbf{Z}$. Our main result shows that $\left(\bigvee_{x\in S} Z(x), \widetilde{\bigvee}_{x\in S} Z(x)\right)$ and $\left(\bigwedge_{x\in S} Z(x), \widetilde{\bigwedge}_{x\in S} Z(x)\right)$ are asymptotically independent. We apply it to nonstationary Gaussian random fields, under appropriate covariance conditions. Conclusions are drawn in Section 4 and the proofs are collected in Appendices.

2. Asymptotic independence of jointly maxima and minima over disjoint discrete subsets of $\mathbb{R}^2$

In the following we show that under an extension of the coordinatewise mixing condition defined in Pereira et al. (2017), the joint maximum and minimum may be regarded as the joint maximum and minimum of an approximately independent sequence of submaxima and subminima.

Throughout the paper we will suppose that there exist sequences of real numbers $\{a_n > 0\}_{n\geq 1}$, $\{c_n > 0\}_{n\geq 1}$, $\{b_n\}_{n\geq 1}$, $\{d_n\}_{n\geq 1}$ and non-degenerate distribution functions $G(x)$ and $H(y)$ such that

$$f(n)\bar{F}(u_n(x)) \underset{n\to+\infty}{\longrightarrow} -\log G(x) \quad f(n)F(v_n(y)) \underset{n\to+\infty}{\longrightarrow} -\log(1 - H(y)), \tag{2.1}$$

where $u_n(x) = a_n x + b_n$ and $v_n(y) = c_n y + d_n$, $x, y \in \mathbb{R}$, and $f(n) \to +\infty$, as $n \to +\infty$.

For each $i = 1, 2$, we shall say that the pair (I, J) is in $\mathcal{S}_i(l_n)$ if $I \subset S$ and $J \subset S$ are subsets of consecutive values of $\pi_i(A_n)$ separated by at least l_n values of $\pi_i(A_n)$, where π_i, $i = 1, 2$, denote the cartesian projections. The cardinality of the sets $\pi_i(A_n)$, $i = 1, 2$, will be denoted by, $\sharp\pi_i(A_n) = f_i(n)$, $i = 1, 2$, and we will assume that $\sharp A_n = f(n) \to +\infty$, as $n \to +\infty$.

Definition 2.1. Let $\mathbf{Z}_A = \{Z(x) : x \in A_n\}_{n\geq 1}$ be a nonstationary sequence and $\{u_n(x_i)\}_{n\geq 1}$ and $\{v_n(y_i)\}_{n\geq 1}$, $i = 1, 2$, $x_1, x_2, y_1, y_2 \in \mathbb{R}$, sequences of real numbers. We say that the sequence $\mathbf{Z}_A$ satisfies condition $D(u_n(x_1), u_n(x_2), v_n(y_1), v_n(y_2))$ if there exist sequences of positive integers $\{l_n\}_{n\geq 1}$ and $\{k_n\}_{n\geq 1}$ such that

$$l_n \underset{n\to+\infty}{\longrightarrow} +\infty, \quad k_n \underset{n\to+\infty}{\longrightarrow} +\infty, \quad k_n l_n \frac{f_i(n)}{f(n)} \underset{n\to+\infty}{\longrightarrow} 0, \ \textit{for each } i = 1, 2, \tag{2.2}$$

and

$$k_n^2 \left[\alpha^*\left(l_n, u_n(x_1), u_n(x_2), v_n(y_1), v_n(y_2)\right) + \alpha_1\left(l_n, u_n(x_1), u_n(x_2)\right) + \alpha_2\left(l_n, v_n(y_1), v_n(y_2)\right)\right] \underset{n\to+\infty}{\longrightarrow} 0,$$

with

1.

$$\alpha^*\left(l_n, u_n(x_1), u_n(x_2), v_n(y_1), v_n(y_2)\right)$$

$$= \sup \left| P\left(\bigcap_{x\in B_1\cup C_1} \{v_n(y_1) < Z(x) \leq u_n(x_1)\}, \bigcap_{x\in B_2\cup C_2} \{v_n(y_2) < Z(x) \leq u_n(x_2)\} \right) - \right.$$