THESIS DECLARATION

The undersigned

Khorrami Chokami, Amir
PhD Registration Number: *1824499*

Thesis Title: *New Advances on Records*

PhD in Statistics
30^{th} Cycle
Advisor: *Professor Simone A. Padoan*
Co-Advisor: *Professor Michael Falk*
Year of Discussion: *2019*

DECLARES

Under *his* responsibility:

1) that, according to the President's decree of 28.12.2000, No. 445, mendacious declarations, falsifying records and the use of false records are punishable under the penal code and special laws, should any of these hypotheses prove true, all benefits included in this declaration and those of the temporary embargo are automatically forfeited from the beginning;

2) that the University has the obligation, according to art. 6, par. 11, Ministerial Decree of 30th April 1999 protocol no. 224/1999, to keep copy of the thesis on deposit at the Biblioteche Nazionali Centrali di Roma e Firenze, where consultation is permitted, unless there is a temporary embargo in order to protect the rights of external bodies and industrial/commercial exploitation of the thesis;

3) that the Servizio Biblioteca Bocconi will file the thesis in its 'Archivio istituzionale ad accesso aperto' and will permit on-line consultation of the complete text (except in cases of a temporary embargo);

4) that in order keep the thesis on file at Biblioteca Bocconi, the University requires that the thesis be delivered by the candidate to Societá NORMADEC (acting on behalf of the University) by online procedure the contents of which must be unalterable and that NORMADEC will indicate in each footnote the following information:

   - thesis *New Advances On Records*;
   - by *Amir Khorrami Chokami*;
   - discussed at Università Commerciale Luigi Bocconi - Milano in *2019*;
   - the thesis is protected by the regulations governing copyright (law of 22 April 1941, no. 633 and successive modifications). The exception is the right of Universitá Commerciale Luigi Bocconi to reproduce the same for research and teaching purposes, quoting the source;

5) that the copy of the thesis deposited with NORMADEC by online procedure is identical to those handed in/sent to the Examiners and to any other copy deposited in the University offices on paper or electronic copy and, as a consequence, the University is absolved from any responsibility regarding errors, inaccuracy or omissions in the contents of the thesis;

6) that the contents and organization of the thesis is an original work carried out by the undersigned and does not in any way compromise the rights of third

parties (law of 22 April 1941, no. 633 and successive integrations and modifications), including those regarding security of personal details; therefore the University is in any case absolved from any responsibility whatsoever, civil, administrative or penal and shall be exempt from any requests or claims from third parties;

7) that the PhD thesis is not the result of work included in the regulations governing industrial property, it was not produced as part of projects financed by public or private bodies with restrictions on the diffusion of the results; it is not subject to patent or protection registrations, and therefore not subject to an embargo.

January 2019

Khorrami Chokami, Amir

Contents

Acknowledgements

I whish to acknowledge Professor Simone Padoan for his precious help, his suggestions, his encouragement and for having always believed in me.

I am grateful to Professor Michael Falk for his illuminating comments, for the wonderful experience as a visiting student at University of Wüerzburg, for his crucial support and contribution in my research activity.

Many thanks to all the members of the Department of Decision Sciences at the Bocconi University, for the extremely constructive atmosphere. Particularly, I would like to thank the other PhD students for being wonderful friends to me.

I am truly grateful to my family, my brother in particular, for inspiring me courage and force of will and for pushing me to do my best.

Milan,
22 January 2019

Amir Khorrami Chokami

Abstract

Record values have always represented a very appealing topic, due to the huge variety of fields they become interesting. For example, breaking a record in the Olympic games can bring an athlete to world fame, while a negative record in stock market indices can lead to a serious economic crisis.

In the univariate case, the commonly employed definition of "record" is the largest random variable in a set of random variables. Under the condition that the position (in time) of a record is know, several results on the stochastic behaviour of records already exists. We focus in the case that position of a record in the sequence of records is unknown. We derive small-sample and large-sample results for the joint distribution of two and three joint records.

Differently from the univariate case, the multivariate theory of records is still in its infancy. Due to the lack of a natural ordering for random vectors, different definitions of a multivariate record are available. We focus on the so-called complete record, that is a random vector whose components are records themselves (each in the univariate sense). We investigate the problem of establishing which results on univariate records do not carry over to the multivariate case. In the case of random vectors with independent components, it is known that the number of records is finite. We compute useful quantities such as the probability that a random vector is the final complete record; furthermore, we find the distribution function of the final complete record. We also derive some results on the arrival times process of complete records.

Finally, we extensively study records over stationary (dependent) gaussian processes. We start by computing the probability that a record takes place and its distribution function. We extend our univariate results computing the joint distributions of two records and of two complete records, respectively. We also provide the asymptotic distribution of a record in the case of a general strictly stationary stochastic process.

List of Papers

The thesis is based on the following papers.

1) Falk, M., **Khorrami, A.** and Padoan, S. A. (2018). *Some Results on Joint Record Events.* Statistics and Probability Letters, 135, 11–19.

2) Falk, M., **Khorrami, A.** and Padoan, S. A. (2018). *On Multivariate Records from Random Vectors withIndependent Components.* Journal of Applied Probability, 55, 1–11.

3) Falk, M., **Khorrami, A.** and Padoan, S. A. (2018). *Records Over Stationary Dependent Sequences.* (arXiv:1807.00337)

Chapter 1

Introduction

1.1 Overview

Extreme value theory consists in creating models to describe risks from rare events with potentially disruptive impacts.

This theory has become a really important and useful discipline over the last decades and its techniques are widely used in many different fields. Many natural phenomena such as precipitation, severe windstorms, flooding, high tides, heavy snow, involve extreme values observed in space.

Air pollution represents one of the most serious problems that modern cities have to face: pollutant concentrations (carbon monoxide, pm10, nitrogen dioxide, etc.) are continuously monitored to trace the frequency and the extent they exceed the thresholds fixed by governments directives.

The heatwave in 2003 was one of the hottest summers Europe experienced in the recent history, an exceptional phenomenon both in terms of intensity and duration.

Other examples of extreme events can be found in the financial world, for example the financial crisis in 2008.

Sports experience positive extreme events, since athletic records are broken in many competitions.

In each circumstance mentioned above, it is of great interest to study the behavior of record values. Suppose we are observing a univariate sequence of values, then the appearance of a record, defined as a value which is greater than the previous ones, implies that the maximum is exceeded. Therefore, studying records can give us information on how often and to which extent do maxima change. Clearly, related questions can concern the expected number of records in a sequence, how to describe the dependence among them, which properties does their arrival times process exhibit, etc.

1.2 Main contributions of the thesis

The theory of records has been extensively developed in the case of sequences of independent and identically distributed random variables, see Chandler (1952) and Arnold, Balakrishnan, and Nagaraja (1998). Also Resnick (1987, Section 4.1), Galambos (1987, Sections 6.2 and 6.3) provide a deep study of records times and record values, showing the connection between records and sample maxima. Let $X_1, X_2, \ldots$ be independent and identically distributed random variables on the real line with a joint continuous distribution function F. The the *record times* are defined as the random indices $\{T(n), n \geq 1\}$ where the maximum changes, i.e. the n-th record appeares at time $T(n)$, while $\{X_{T(n)}, n \geq 1\}$ are called the *record values*. Galambos (1987) proves the Markovian structure of the arrival times process of records, as well as the independence of the Bernoullian random variables describing the occurrence of records and the asymptotic distribution of $X_{T(n)}$ as $n \to \infty$. Resnick (1987) focuses on the record value process, showing that it is a Poisson process, and also the arrival times process is asymptotically Poisson, if the underlying distribution is continuous.

Alternatively to that, we investigate the stochastic behavior of arbitrary X_j, X_k, $j < k$, under the condition that they are records, without knowing their orders in the sequence of records. The results are completely different. In particular it turns out that the distribution of X_k, being a record, is not affected by the additional knowledge that X_j is a record as well. On the contrary, the distribution of X_j, being a record, is affected by the additional knowledge that X_k is a record as well.

If F has a density, then the gain of this additional information, measured by the corresponding Kullback-Leibler distance, is j/k, independent of F. We derive the limiting joint distribution of two records, which is not a bivariate extreme value distribution. We extend this result to the case of three records. In a special case we also derive the limiting joint distribution of increments among records. This is the content of Chapter 2.

Records can also be studied over sequences of random vectors, although the lack of a natural ordering in $\mathbb{R}^d$ leads to various definitions of records and some of them are listed in Goldie and Resnick (1989). In Chapter 3 we deal with the so called *complete records*. Let $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots$ be independent copies of a random vector $\boldsymbol{X}$ with values in $\mathbb{R}^d$ and with a continuous distribution function. The random vector $\boldsymbol{X}_n$ is a complete record, if each of its components is a record. As we require $\boldsymbol{X}$ to have independent components, crucial results for univariate records clearly carry over. But there are substantial differences as well: While there are infinitely many records in case $d = 1$, there occur only finitely many in the series if $d \geq 2$. In fact, the probability of occurrence of a complete record decreases too rapidly as time passes by, so to prevent the complete record to appear an infinite number of times. Consequently, there is a terminal complete record with probability one. We compute the distribution of the random total number of complete records and investigate the distribution of the terminal record. For complete records, the sequence of waiting times forms a Markov chain, but differently from the univariate case, now the state infinity is an absorbing element of the state space.

The analysis of records presented so far is based on a very strong assumption, which is not satisfied in practical applications: the independence in the underlying sequence of random values/vectors. In fact, there are few results in the literature on the behavior of records and record times for other types of sequences of random variables or random vectors. Ballerini and Resnick (1987) consider records over random sequences with a linear trend, while Arnold, Balakrishnan, and Nagaraja (1998) deal with the case of records over Poisson processes, birth processes, and renewal processes.

In Chapter 4 we derive the probability that a record at the time n, say X_n, takes place over a zero-mean, unit-variance second-order stationary univariate Gaussian process. We also compute its distribution function. We study the joint distribution of the arrival time process of records and the distribution of the increments between the first and second record and the third and second record and we compute the expected number of records. We also consider two consecutive and non-consecutive records, one at time j and one at time n and we derived the probability that the joint record (X_j, X_n) takes place and its distribution function. The probability that the records X_n and (X_j, X_n) take place and the arrival time of the n-th record, are independent of the marginal distribution function, provided that is continuous. These results actually hold for a second-order stationary process with Gaussian copulas. We extend some of these results to the case of multivariate Gaussian process. Finally, for a strictly stationary process satisfying some mild conditions on the tail behavior of the common marginal distribution function F and the long-range dependence of the extremes of the process we derive the asymptotic probability that the record X_n takes place and its distribution function.

1.3 Preliminary Background

1.3.1 Extremes of independent univariate sequences

In this section we provide a brief description of the extremal behavior of independent and identically distributed (i.i.d.) random sequences. Here and throughout the thesis, the converge in distribution is indicated with the arrow $\rightarrow$.

Let $X_1, X_2, \ldots$ be a sequence of i.i.d. random variables following a continuous distribution function F. It is easy to see that the distribution of the n-partial maximum $M_n = \max(X_1, \ldots, X_n)$ degenerates to a point mass on the upper endpoint of F. In order to obtain a non-degenerate distribution, it is enough to suitably rescale M_n, as shown in the following theorem

Theorem 1.3.1. *If there exist sequences of constants $\{a_n > 0\}$ and $\{b_n \in \mathbb{R}\}$ such that*

$$\Pr\{(M_n - b_n)/a_n \leq z\} \rightarrow G(x) \qquad n \rightarrow \infty \tag{1.1}$$

for a non-degenerate distribution function G, then G must be a member of the generalized extreme value (GEV) family of distributions $G_\gamma(ax+b)$, for $a > 0, b \in \mathbb{R}$, i.e.

$$G_\gamma(x) = \exp\left\{-[1+\gamma z]^{-\frac{1}{\gamma}}\right\}, \quad 1+\gamma x > 0$$

where $\gamma \in \mathbb{R}$.

The cdfs of the three sub-classes of the GEV, i.e. the Gumbel, Fréchet, and negative Weibull are denoted by $G_0(x) = \exp\left(\mathrm{e}^{-x}\right)$, $G_\alpha(x) = G_{1/\gamma}((x-1)/\gamma)$ for $\gamma > 0$ and $G_\beta(x) = G_{-1/\gamma}(-(x+1)/\gamma)$ for $\gamma < 0$ (see e.g., Falk, Hüsler, and Reiss 2011, Ch. 2 for details).

1.3.2 Multivariate models

Let (X_1, Y_1), $(X_2, Y_2)\ldots$ be a sequence of i.i.d. random vectors with distribution function $F(x, y)$. In order to characterize the extremal behavior of multivariate extremes, we adopt an approach which is simlar to that used in the univariate case.

We define the *vector of componentwise maxima* as follows:

$$\boldsymbol{M}_n = (M_{x,n}, M_{y,n}) := (\max_{i=1,\dots,n} X_i, \max_{j=1,\dots,n} Y_j).$$

Note that $\boldsymbol{M}_n$ need not to be an observed vector of the original series.

As in the univariate case, we are interested to study the distribution function of $\boldsymbol{M}_n$ as $n \to \infty$. We can apply Theorem 1.3.1 to the marginal series $\{X_i\}$ and $\{Y_i\}$ separately, also by assuming a known marginal distribution for both the X_i and the Y_i. The most used representation assumes this marginal univariate distribution to be a standard Fréchet distribution:

$$F(z) = \exp\left(-\frac{1}{z}\right), \qquad z > 0.$$

This procedure leads to the parametric part of the model. The non parametric part arises in the description of the dependence between the extremes. By renormalizing $\boldsymbol{M}_n$, by taking $a_n = n$ and $b_n = 0$ we obtain

$$\boldsymbol{M}_n^* = \left(\max_{i=1,\dots,n} \{X_i\}/n, \max_{i=1,\dots,n} \{Y_i\}/n\right), \tag{1.2}$$

We are now ready to state the following theorem (Coles et al. (2001, Ch. 8))

Theorem 1.3.2. *If* $\boldsymbol{M}_n^* = (M_{x,n}^*, M_{y,n}^*)$ *satisfies*

$$\Pr(M_{x,n}^* \leq x, M_{y,n}^* \leq y) \to G(x, y),$$

for a non-degenerate distribution function G*, then* G *belongs to the family of the* bivariate extreme value distributions

$$G(x, y) = \exp\{-V(x, y)\} \tag{1.3}$$

with $x > 0, y > 0$*,*

$$V(x, y) = 2\int_0^1 \max\left(\frac{w}{x}, \frac{1-w}{y}\right) dH(w), \tag{1.4}$$

and H *is a distribution function on* $[0, 1]$ *such that*

$$\int_0^1 w \, dH(w) = \frac{1}{2}.$$

The function V is called the *exponent measure* and has the following properties:

(i)
$$V(x,\infty) = \frac{1}{x}, \quad V(\infty,y) = \frac{1}{y},$$

(ii)
$$V(ax,ay) = a^{-1}V(x,y), \quad a > 0.$$

This results can be extended to the case of d-dimensional random vectors, provided that the corresponding exponent function $V(z_1,\ldots,z_d)$ satisfies the properties previously stated.

1.3.3 Extremes over dependent sequences

The results mentioned so far, although mathematically elegant, are based on the assumption of time independent rvs, which poorly models real world processes where data come from. Dealing with extremes of a general stochastic processes is, however, still an unsolvable problem, unless we impose some constraints on the dependence structure of the process.

An extension of the results introduced in the previous section can be obtained focusing on a stationary sequence of dependent but identically distributed random variables.

Let $\{X_n, n \geq 1\}$ be a strictly stationary sequence of rvs, i.e. the joint distribution of $(X_{j_1},\ldots,X_{j_n})$ and $(X_{j_1+m},\ldots,X_{j_n+m})$ are identical, for every n, m and $j_1,\ldots,j_n$. Precisely, we assume that F belongs to the (maximum) domain of attraction of G_γ, in symbols $F \in \mathcal{D}(G_\gamma)$, $\gamma \in \mathbb{R}$. This means that (1.1) holds for the sequence of i.i.d. rvs $Y_1,\ldots,Y_n$ with common cdf F. Concerning the dependence structure of $\{X_n, n \geq 1\}$ we assume a mild condition on the long-range dependence of extremes of such a stationary sequence, see Leadbetter et al. (1983, Ch. 3) for details. Partition $\{1,\ldots,n\}$ into $k_n = \lfloor n/r_n \rfloor$ blocks of length $r_n = o(n)$. Suppose that for every $\lambda > 0$, there is a sequence of real-value thresholds $u_n(\lambda)$, $n = 1, 2\ldots,$

such that

$$\lim_{n\to\infty} n\,\Pr(X_1 > u_n(\lambda)) = \lambda, \quad \lambda > 0$$

and the condition $D(u_n(\lambda))$ is satisfied for each such λ. Specifically, for every $\lambda > 0$, let

$$K_n(l) = \max(|\Pr(X_i \le u_n(\lambda), i \in I \cup J) - \Pr(X_i \le u_n(\lambda), i \in I)\Pr(X_i \le u_n(\lambda), i \in J)|)$$

where $I, J \subset \{1, \dots, n\}$ such that $\min\{|i-j| : i \in I, j \in J\} = l$. Then, we say that condition $D(u_n(\lambda))$ holds for each such λ, if $K_n(l_n) \to 0$ as $n \to \infty$ for some sequence $l_n \to \infty$ with $l_n = o(n)$ (pp. 53-57, Leadbetter et al. 1983). By Lemma 3.2.2 in Leadbetter et al. (1983) this means that extreme events, such as the partial maxima $M_{E_i} = \max_{j\in E_i}(X_j)$, with $E_i = \{(i-1)r_n+1, \dots, ir_n\} \setminus \{ir_n - l_n + 1, \dots, ir_n\}$, $i = 1, \dots, k_n$, which are separated by l_n, are almost independent.

Then, under these conditions by Leadbetter et al. (1983, Theorem 3.7.1) we have that for suitable choices of $r_n \to \infty$ with $n \to \infty$ such that $k_n K_n(l_n) \to 0$ and $k_n l_n \to 0$ as $n \to \infty$, it follows that

$$\lim_{n\to\infty} \Pr(\max(X_1, \dots, X_n) \le u_n(\lambda)) = \exp(-\theta\lambda), \quad 0 < \theta \le 1,\ \lambda > 0. \tag{1.5}$$

When this holds true we say that the sequence $\{X_n, n \ge 1\}$ has *extremal index* $\theta \in (0, 1]$. The result in (1.5) implies that for $\lambda = -\log G_\gamma(x)$, $x \in \mathbb{R}$, and suitable norming constants $a_n > 0$ and $b_n \in \mathbb{R}$,

$$\lim_{n\to\infty} \Pr\left(\max(X_1, \dots, X_n) \le a_n x + b_n\right) = G_\gamma(x)^\theta, \quad x, \gamma \in \mathbb{R},$$

Loosely speaking, the extremal index is a parameter that quantifies the impact that the dependence structure of the stationary sequence has on the asymptotic distribution of extreme events such as the partial maximum M_n, for sufficiently large n. When $\theta = 1$ we recover (1.1), i.e. the asymptotic distribution of the normalized maximum for a sequence of independent variables. When $\theta < 1$, then for every $x \in \mathbb{R}$ we have that $G_\gamma(x) \le G_\gamma^\theta(x)$ and therefore $1 - G_\gamma(x) \ge 1 - G_\gamma^\theta(x)$. In other words, the dependence of the stationary sequence reduces the size of the extreme events.

1.4 Records over univariate i.i.d. sequences of random variables

Let $X_1, X_2, \dots$ be independent and identically distributed (i.i.d.) random variables (rvs), with a continuous distribution function F. The rv X_m is a record if $X_m > \max(X_1, \dots, X_{m-1})$. Clearly, X_1 is a record.

Records in this framework have been extensively investigated over the past decades, with useful results on record times and limit laws for record values. This section reviewes this background by highlighting the most important results, with the scope of a better understanding of the novelties presented in the following chapters. Topics covered in the following lines can be found in Resnick (1987, Section 4.1), Galambos (1987, Sections 6.2 and 6.3), and Arnold, Balakrishnan, and Nagaraja (1998).

In the framework of i.i.d. univariate sequences, we define the *record times* as follows: put $T(1) := 1$ and, for $n \geq 2$,

$$T(n) := \min\{m > n-1 \colon X_m \text{ is a record}\}.$$

Then, $X_{T(n)}$, $n \in \mathbb{N}$, is the sequence of records among $X_1, X_2, \dots$ and $T(n)$ is the arrival time of the $n-th$ record. The most useful property concerning record times lies in the fact that the record time process is a Markov chain. Precisely, let us firstly state the following lemma.

Lemma 1.4.1. *For any integers* $1 = j_1 < j_2 < \cdots < j_n$, $n = 1, 2, \dots$, *it holds that*

$$\Pr\left(T(i) = j_i,\ 1 \leq i \leq n\right) = \frac{1}{j_n \prod_{i=2}^{n-1}(j_i - 1)}.$$

From the previous lemma, it can clearly be seen that the following theorem holds true.

Theorem 1.4.2. *Let* $X_1, X_2, \dots$ *be an i.i.d. sequence of random variables. Then the record times process is a Markov chain with transitions*

$$\Pr\left(T(n+1) = m | T(n) = k\right) = \frac{k}{m(m-1)}, \quad 1 \leq k < m.$$

Consider now the indicator function

$$I_m := \mathbb{1}(X_m \text{ is a record}), \qquad m \in \mathbb{N}.$$

It is well known that the indicator functions $I_1, I_2, \dots$ are independent with

$$\Pr(I_m = 1) = m^{-1},\ m \in \mathbb{N}, \tag{1.6}$$

see, e.g., Galambos (1987, Lemma 6.3.3). An immediate consequence of the previous equation is that

$$\Pr(I_1 = 1, \dots I_m = 1) = (m!)^{-1}, \qquad m \in \mathbb{N}.$$

Record indicators are particularly useful when we deal with the number of records in a sequence. In fact, an obvious way to describe the number of records in the set $X_1, \dots, X_n$ is

$$N_n = \sum_{j=1}^{n} I_j, \qquad n \in \mathbb{N},$$

and it is evident that

$$\mathrm{E}[N_n] = \sum_{j=1}^{n} j^{-1} \xrightarrow[n\to\infty]{} \infty.$$

Suppose now that the common distribution function (df) F of $X_1, X_2, \dots$ is the standard exponential df $F(x) = 1-\exp(-x)$, $x \geq 0$. The increments of subsequent records are given by the sequence

$$Y_n := X_{T(n)} - X_{T(n-1)},\ n \geq 2, \quad Y_1 := X_1.$$

Then, $Y_1, Y_2, \dots$ are i.i.d. rvs with common df $F(x) = 1 - \exp(-x)$, $x \geq 0$ (see Galambos (1987) for the proof). This yields

$$X_{T(n)} = \sum_{i=1}^{n} Y_i,\ n \in \mathbb{N}, \tag{1.7}$$

and, thus, characterizes the distribution of the n-th record or the joint distribution of several numbered records $\left(X_{T(n_1)}, X_{T(n_2)}, \dots, X_{T(n_m)}\right)$, $n_1 < n_2 < \cdots < n_m$, etc.

But what can be said for a general sequence $X_1, X_2, \dots$ of i.i.d. rvs? The following theorem provides an answer to this question by finding the class of limiting distribution functions for record values (see Resnick (1973) for details).

Theorem 1.4.3. *Let $X_1, X_2, \ldots$ be an i.i.d. sequence of random variables with continuous distribution F. Then, all possible limit distribution functions H such that*

$$\Pr(X_{T(n)} \leq A_n x + B_n) \underset{n\to\infty}{\longrightarrow} H(x) \tag{1.8}$$

have the form

$$H(x) = \Phi(g(x)),$$

where Φ is the cdf of the standard normal distribution and $g(x)$ can be one of the following functions

(i) $g(x) = x,$

(ii) $g(x) = \alpha \log x,$ *for* $\alpha > 0,$ *if* $x > 0,$

(iii) $g(x) = -\alpha \log(-x),$ *for* $\alpha > 0,$ *if* $x < 0.$

Chapter 2

Some Results on Joint Records

2.1 Introduction

The results on records presented in the previous chapter are based on the fact that we know which record we are considering. Differently to that, in this chapter we drop the assumption that we *know* the order of a record. Therefore, we characterize the distribution

$$\Pr\left(X_j \leq \cdot \mid X_j \text{ is a record}\right),\ j \in \mathbb{N},$$

as well as the joint distribution of two records

$$\Pr\left(X_j \leq \cdot, X_k \leq \cdot \mid X_j \text{ and } X_k \text{ are records}\right),\ 1 \leq j < k.$$

We achieve this under the assumption that the joint df F of $X_1, X_2, \ldots$ is continuous. In particular, we establish the following surprising fact: Choose integers $j < k$. The distribution of X_j, being a record, is affected when we know that X_k is a record as well. The distribution of X_k, being a record, however, is not affected when we know that X_j is a record as well. The corresponding information gain is measured by the Kullback-Leibler distance between the densities. This information gain turns out to be j/k and it is independent of the underlying F. This is the content of Section 1.

In Section 2, the asymptotic joint distribution of X_j and X_k, suitably standardized, under the condition that they are records, is derived. This is achieved if the underlying df F is in the domain of attraction of an extreme value df. The limit distribution is not an extreme value distribution. We also derive the limiting joint distribution of three records. Finally, for the special case of a sequence of iid rvs with a common standard negative exponential distribution, we derive the asymptotic joint distribution of increments among records.

2.2 Distribution of Records

Throughout this section we suppose that $X_1, X_2 \dots$ are iid rvs with a common continuous df F. The distribution of X_n, being a record, is provided by the following important result.

Lemma 2.2.1. *We have for* $n \in \mathbb{N}$

$$\Pr\left(X_n \le x \mid X_n \text{ is a record}\right) = \Pr\left(\max_{1\le i\le n} X_i \le x\right) = F^n(x),\ x \in \mathbb{R}.$$

Therefore, the distribution of X_n, being a record, coincides with that of the largest order statistic in the sample $X_1, \dots, X_n$.

Proof. Denote by $X_{1:n} \le \dots \le X_{n:n}$ the order statistics pertaining to $X_1, \dots, X_n$, and by $R(X_i) = \sum_{j=1}^n \mathbb{1}(X_j \le X_i)$ the rank of X_i, $1 \le i \le n$. It is well known that the vector of order statistics $(X_{1:n}, \dots, X_{n:n})$ and the vector of ranks $(R(X_1), \dots, R(X_n))$ are independent, with $\Pr(R(X_i) = k) = n^{-1}$, $1 \le i, k \le n$; see, e.g., Rényi (1962). Therefore, we obtain from equation (1.6),

$$\begin{aligned}
&\Pr\left(X_n \le x \mid X_n \text{ is a record}\right) = n \Pr\left(X_n \le x, X_n \text{ is a record}\right) \\
&= n \Pr\left(X_{n:n} \le x, R(X_n) = n\right) = \Pr\left(X_{n:n} \le x\right),
\end{aligned}$$

which is the assertion. □

The above result immediately yields the limiting distribution of X_n, being a record, as n tends to infinity. The necessary tools are provided by univariate extreme value theory: Suppose that there exist constants $a_n > 0, b_n \in \mathbb{R}, n \in \mathbb{N}$, such that

$$F^n\left(a_n x + b_n\right) \underset{n\to\infty}{\longrightarrow} G(x), \quad x \in \mathbb{R},$$

for all continuity point x of G, where G is a non-degenerate df. Then, F is said to be in the *max-domain of attraction* of G, denoted by $F \in \mathcal{D}(G)$, and G is a univariate extreme value distribution. Precisely, G is a member of a parametric family $\{G_\alpha \colon \alpha \in \mathbb{R}\}$, indexed by $\alpha \in \mathbb{R}$, with

$$G_\alpha(x) = \exp\left(-(1+\alpha x)^{-1/\alpha}\right), \quad 1+\alpha x > 0,$$

if α is different from zero, and the convention

$$G_0(x) = \lim_{\alpha \to 0} G_\alpha(x) = \exp\left(-\mathrm{e}^{-x}\right), \quad x \in \mathbb{R},$$

(see, e.g., Resnick 1987 Ch. 1). If we put in particular $F(x) = 1 - \exp(-x)$, $x \geq 0$, then we have $F \in \mathcal{D}(G_0)$, precisely

$$F^n(x + \log n) \xrightarrow[n\to\infty]{} \exp\left(-\mathrm{e}^{-x}\right), \quad x \in \mathbb{R},$$

and, thus,

$$\Pr\left(X_n - \log n \leq x \mid X_n \text{ is a record}\right) \xrightarrow[n\to\infty]{} \exp\left(-\mathrm{e}^{-x}\right), \quad x \in \mathbb{R}.$$

Next we establish the joint distribution of two records. To simplify the notation, we suppose that the underlying df of the sequence of iid rvs is the standard negative exponential df $F(x) = \exp(x)$, $x \leq 0$. Instead of writing $X_1, X_2, \dots$ we use with this particular underlying df the notation $\eta_1, \eta_2, \dots$. The latter distribution is a member of the set $\{G_\alpha \colon \alpha \in \mathbb{R}\}$, with $\alpha = -1$ and shifted by -1. In this particular case we have $F^n(x) = \exp(nx) = F(nx)$, $x \leq 0$.

Lemma 2.2.2. *We have for* $1 \leq j < k$ *and* $x_1, x_2 \in \mathbb{R}$*,*

$$\begin{aligned}
&\Pr(\eta_j \leq x_1, \eta_k \leq x_2 \mid \eta_j \textit{ and } \eta_k \textit{ are records}) \\
&= \frac{k}{k-j} \Pr\left(\eta_1 \leq j x_1, \eta_2 \leq (k-j)x_2, (k-j)\eta_1 < j\eta_2\right) \\
&= \begin{cases} \frac{k}{k-j}\left(\mathrm{e}^{(k-j)x_2}\,\mathrm{e}^{jx_1} - \frac{j}{k}\,\mathrm{e}^{kx_1}\right), & \textit{if } x_1 < x_2 \\ \mathrm{e}^{kx_2} = \Pr(\eta_k \leq x_2 \mid \eta_k \textit{ is a record}), & \textit{if } x_1 \geq x_2. \end{cases}
\end{aligned}$$

Proof. Let $\eta_1^{(r)}, \eta_2^{(r)}, \dots,\ r = 1, 2$, be two independent sequences of iid copies of η. Let $\eta_{i:n}^{(r)}$ be i-th order statistics and $R_m^{(r)}(\eta_j^{(r)})$ the rank of $\eta_j^{(r)}$ in the sample $\eta_1^{(r)}, \dots, \eta_m^{(r)}$. We split the sample $\eta_1, \dots, \eta_k$ into the two independent sub-samples $(\eta_1, \dots, \eta_j) =: (\eta_1^{(1)}, \dots, \eta_j^{(1)})$ and $(\eta_{j+1}, \dots, \eta_k) =: (\eta_1^{(2)}, \dots, \eta_{k-j}^{(2)})$. By the independence between vectors of order statistics and ranks and the fact that the distributions of $\eta_{m:m}$ and η/m coincide for $m \in \mathbb{N}$, we obtain

$$\begin{aligned}
&\Pr(\eta_j \le x_1, \eta_k \le x_2 \mid \eta_j \text{ and } \eta_k \text{ are records})\\
&= jk \Pr\Big(\eta_{j:j}^{(1)} \le x_1,\ \eta_{k-j:k-j}^{(2)} \le x_2,\ R_j^{(1)}(\eta_j^{(1)}) = j,\\
&\qquad\qquad R_{k-j}^{(2)}(\eta_{k-j}^{(2)}) = k-j,\ \eta_{j:j}^{(1)} < \eta_{k-j:k-j}^{(2)}\Big)\\
&= jk \Pr\Big(\eta_{j:j}^{(1)} \le x_1,\ \eta_{k-j:k-j}^{(2)} \le x_2,\ \eta_{j:j}^{(1)} < \eta_{k-j:k-j}^{(2)}\Big)\\
&\qquad \times \Pr\Big(R_j^{(1)}(\eta_j^{(1)}) = j\Big) \Pr\Big(R_{k-j}^{(2)}(\eta_{k-j}^{(2)}) = k-j\Big)\\
&= \frac{k}{k-j} \Pr(\eta_1 \le jx_1,\ \eta_2 \le (k-j)x_2,\ (k-j)\eta_1 < j\eta_2).
\end{aligned}$$

The rest of the assertion follows from elementary computations, conditioning on η_2. □

The preceding result can be extended to $X_1, X_2, \dots$ with an arbitrary continuous df F by putting $X_i := F^{-1}(\exp(\eta_i))$, $i \in \mathbb{N}$, where $F^{-1}(q) := \inf\{t \in \mathbb{R} \colon F(t) \ge q\}$, $q \in (0,1)$, is the usual generalized inverse of F. From the general equivalence $F^{-1}(q) \le t$ iff $q \le F(t)$, $q \in (0,1)$, $t \in \mathbb{R}$, we obtain for $1 \le j < k$ and $y_1, y_2 \in \mathbb{R}$

$$\begin{aligned}
&\Pr(X_j \le y_1, X_k \le y_2 \mid X_j \text{ and } X_k \text{ are records})\\
&= \Pr(\eta_j \le \log(F(y_1)), \eta_k \le \log(F(y_2)) \mid \eta_j \text{ and } \eta_k \text{ are records}).
\end{aligned}$$

By putting $x_i := \log(F(y_i))$, $i = 1, 2$, the following general result is an immediate consequence of Lemma 2.2.2.

Corollary 2.2.3. *We have for integers $1 \le j < k$ and $y_1, y_2 \in \mathbb{R}$*

$$\begin{aligned}
&\Pr(X_j \le y_1, X_k \le y_2 \mid X_j \text{ and } X_k \text{ are records})\\
&= \begin{cases} F^j(y_1)\left(\frac{k}{k-j} F^{(k-j)}(y_2) - \frac{j}{k-j} F^{k-j}(y_1)\right), & \text{if } F(y_1) < F(y_2)\\ F^k(y_2) = \Pr(X_k \le y_2 \mid X_k \text{ is a record}), & \text{if } F(y_2) \le F(y_1). \end{cases}
\end{aligned}$$

Choose integers $1 \le j < k$. Next we establish the fact that the distribution of η_j, being a record, is affected, if we know that η_k is a record as well. The distribution of η_k, being a record, however, is not affected by the additional knowledge that η_j is a record as well.

Proposition 2.2.4. *We have for integers $1 \le j < k$ and $x_1, x_2 \in \mathbb{R}$,*

$$\begin{aligned}\Pr\left(\eta_k \le x_2 \mid \eta_j \textit{ and } \eta_k \textit{ are records}\right) &= \mathrm{e}^{kx_2}, \qquad x_2 \le 0,\\ &= \Pr(\eta_k \le x_2 \mid \eta_k \textit{ is a record})\end{aligned}$$

and

$$\begin{aligned}&\Pr\left(\eta_j \le x_1 \mid \eta_j \textit{ and } \eta_k \textit{ are records}\right)\\ &= \frac{k}{k-j}\,\mathrm{e}^{jx_1} - \frac{j}{k-j}\,\mathrm{e}^{kx_2} \qquad x_1 \le 0,\\ &= \frac{k}{k-j}\Pr(\eta_j \le x_1 \mid \eta_j \textit{ is a record}) - \frac{j}{k-j}\Pr(\eta_k \le x_2 \mid \eta_k \textit{ is a record}).\end{aligned}$$

Proof. From Lemma 2.2.2 we obtain

$$\Pr(\eta_k \le x_2 \mid \eta_j \text{ and } \eta_k \text{ are records}) = \frac{k}{k-j}\Pr\left(\eta_2 \le (k-j)x_2, (k-j)\eta_1 < j\eta_2\right)$$

by putting $x_1 = 0$, and

$$\Pr(\eta_j \le x_1 \mid \eta_j \text{ and } \eta_k \text{ are records}) = \frac{k}{k-j}\Pr\left(\eta_1 \le jx_1, (k-j)\eta_1 < j\eta_2\right)$$

by putting $x_2 = 0$. The assertion is now an immediate consequence of elementary computations by conditioning on η_2. □

The df

$$F_{j,k}(x) = \frac{k}{k-j}\,\mathrm{e}^{jx} - \frac{j}{k-j}\,\mathrm{e}^{kx}, \qquad x \le 0,$$

has the density

$$f_{j,k}(x) = \frac{jk}{k-j}\left(\mathrm{e}^{jx} - \mathrm{e}^{kx}\right), \qquad x \le 0.$$

Suppose η_j is a record. To summarize by a single number the information, which is inherent in the additional knowledge that η_k is a record as well, we compute the Kullback-Leibler divergence between the density $f_{j,k}$ and the density $f_j(x) =$

$j\exp(jx)$, which is the density of the df $F_j(x) = \Pr(\eta_j \le x \mid \eta_j \text{ is a record}) = \exp(jx)$, $x \le 0$.

In a general context, the Kullback-Leibler divergence of a density $q(\cdot)$ from a density $p(\cdot)$, is defined by

$$D_{KL}(p||q) := \int_{-\infty}^{+\infty} p(x) \log \frac{p(x)}{q(x)} \mathrm{d}x.$$

It quantifies the information lost, when $q(\cdot)$ is used to approximate $p(\cdot)$. Closely related to the Kullback-Leibler divergence, the Kullback-Leibler distance of p and q is defined by

$$D_{KL}(p, q) := D_{KL}(p||q) + D_{KL}(q||p).$$

Note that $D_{KL}(p||q) \ge 0$ by Jensen's inequality.

Proposition 2.2.5. *The Kullback-Leibler distance between the densities $f_{j,k}$ and $f_j(x) = j\exp(x)$, $x \le 0$, is given by*

$$D_{KL}(f_{j,k}, f_j) = j/k, \qquad k > j \ge 1.$$

Clearly, $0 < D_{KL}(f_{j,k}, f_j) < 1$. The Kullback-Leibler distance between the densities $f_{j,k}$ and f_j gets small if j/k gets small. This means that the additional knowledge that η_k is a record as well, affects the distribution of η_j, being a record, less if k gets large. On the other hand, if $k = j + 1$, then the information gain approaches one if j gets large.

Proof. We have

$$\begin{aligned}
D_{KL}(f_{j,k} \| f_j) &= \int_{-\infty}^{0} \frac{jk}{k-j} \left(\mathrm{e}^{jx} - \mathrm{e}^{kx}\right) \log \left(\frac{jk}{k-j} \left(\mathrm{e}^{jx} - \mathrm{e}^{kx}\right) \frac{\mathrm{e}^{-jx}}{j} \right) \mathrm{d}x \\
&= \frac{jk}{k-j} \left(\int_{-\infty}^{0} \left(\mathrm{e}^{jx} - \mathrm{e}^{kx}\right) \log \frac{k}{k-j} \mathrm{d}x \right. \\
&\qquad \left. + \int_{-\infty}^{0} \left(\mathrm{e}^{jx} - \mathrm{e}^{kx}\right) \log \left(1 - \mathrm{e}^{(k-j)x}\right) \mathrm{d}x \right) \\
&= \log \frac{k}{k-j} + \frac{jk}{k-j} \int_{-\infty}^{0} \left(\mathrm{e}^{jx} - \mathrm{e}^{kx}\right) \log \left(1 - \mathrm{e}^{(k-j)x}\right) \mathrm{d}x.
\end{aligned}$$

The substitution $t = 1 - \mathrm{e}^{(k-j)x}$ entails

$$\int_{-\infty}^{0} \left(\mathrm{e}^{jx} - \mathrm{e}^{kx}\right) \log\left(1 - \mathrm{e}^{(k-j)x}\right) \mathrm{d}x = \frac{1}{k-j} \int_0^1 \Big((1-t)^{\frac{j}{k-j}-1} - (1-t)^{\frac{k}{k-j}-1} \Big) \log t \mathrm{d}t$$

with

$$\begin{aligned}\int_0^1 (1-t)^{\frac{j}{k-j}-1} \log t \mathrm{d}t &= B\left(1, \frac{j}{k-j}\right)\left(\psi(1) - \psi\left(1 + \frac{j}{k-j}\right)\right) \\ &= \frac{k-j}{j}\left(\psi(1) - \psi\left(1 + \frac{j}{k-j}\right)\right),\end{aligned}$$

where $B(a,b) = \int_0^1 t^{a-1}(1-t)^{b-1}\mathrm{d}t$, $a, b > 0$, is the Beta function and $\psi(x) = \Gamma'(x)/\Gamma(x)$, $x > 0$ is the Digamma function; $\Gamma(x) = \int_0^\infty t^{x-1}\,\mathrm{e}^{-t}\,\mathrm{d}t$, $x > 0$, is the usual Gamma function. In an analogous way, one obtains

$$\int_0^1 (1-t)^{\frac{k}{k-j}-1} \log t \mathrm{d}t = \frac{k-j}{k}\left(\psi(1) - \psi\left(1 + \frac{k}{k-j}\right)\right).$$

As a consequence, we obtain

$$D_{KL}(f_{j,k} \| f_j) = \log\frac{k}{k-j} + \psi(1) + \frac{j}{k-j}\psi\left(1 + \frac{k}{k-j}\right) - \frac{k}{k-j}\psi\left(1 + \frac{j}{k-j}\right).$$

Furthermore, we have,

$$\begin{aligned}D_{KL}(f_j \| f_{j,k}) &= -\int_{-\infty}^{0} j\,\mathrm{e}^{jx} \log\left(\frac{jk}{k-j}\left(\mathrm{e}^{jx} - \mathrm{e}^{kx}\right)\frac{\mathrm{e}^{-jx}}{j}\right)\mathrm{d}x \\ &= -\log\frac{k}{k-j}\int_{-\infty}^{0} j\,\mathrm{e}^{jx}\,\mathrm{d}x - j\int_{-\infty}^{0} \mathrm{e}^{jx}\log\left(1 - \mathrm{e}^{(k-j)x}\right)\mathrm{d}x \\ &= -\log\frac{k}{k-j} - \frac{j}{k-j}\int_0^1 (1-t)^{\frac{j}{k-j}-1}\log t \mathrm{d}t \\ &= -\log\frac{k}{k-j} - \frac{j}{k-j}\frac{k-j}{k}\left(\psi(1) - \psi\left(1 + \frac{k}{k-j}\right)\right) \\ &= -\log\frac{k}{k-j} - \psi(1) + \psi\left(1 + \frac{k}{k-j}\right).\end{aligned}$$

Finally, we obtain

$$\begin{aligned}
D_{KL}(f_{j,k}\|f_j) + D_{KL}(f_j\|f_{j,k}) &= \frac{j}{k-j}\psi\left(1+\frac{k}{k-j}\right) \\
&\quad + \left(1-\frac{k}{k-j}\right)\psi\left(1+\frac{j}{k-j}\right) \\
&= \frac{j}{k-j}\psi\left(1+\frac{k}{k-j}\right) - \frac{j}{k-j}\psi\left(1+\frac{j}{k-j}\right) \\
&= \frac{j}{k-j}\left(\psi\left(1+\frac{k}{k-j}\right) - \psi\left(1+\frac{j}{k-j}\right)\right).
\end{aligned}$$

The functional equation $\psi(1+x) = \psi(x) + 1/x$, $x > 0$, implies

$$\psi\left(1+\frac{k}{k-j}\right) - \psi\left(1+\frac{j}{k-j}\right) = \psi\left(1+\frac{k}{k-j}\right) - \psi\left(\frac{k}{k-j}\right) = \frac{k-j}{j},$$

which yields the assertion. □

It turns out that the preceding result carries over to records from a sequence $X_1, X_2, \dots$ of iid rvs with an arbitrary df F that has a density, say f. Choose integers $1 \le j < k$. From Corollary 2.2.3 and Proposition 2.2.4 we obtain that the density function of the df

$$G_{j,k}(x) := \Pr(X_j \le x \mid X_j \text{ and } X_k \text{ are records})$$

is

$$g_{j,k}(x) = \frac{jk}{k-j} f(x)\left(F^{j-1}(x) - F^{k-1}(x)\right), \qquad x \in \mathbb{R},$$

and the density function of the df

$$G_j(x) := \Pr(X_j \le x \mid X_j \text{ is a record})$$

is

$$g_j(x) = j f(x) F^{j-1}(x), \qquad x \in \mathbb{R}.$$

Proposition 2.2.6. *The Kullback-Leibler distance between the densities $g_{j,k}$ and g_j is given by*

$$D_{KL}(g_{j,k}, g_j) = j/k,$$

for $k > j \ge 1$.

The preceding result shows that j/k is a universal constant, valid for each df F that has a density, which reflects the gain of information contained in the additional knowledge that X_k is a record as well.

Proof. We have

$$D_{KL}(g_{j,k}\|g_j) = \int_{-\infty}^{+\infty} \frac{jk}{k-j} f(x)\left(F^{j-1}(x) - F^{k-1}(x)\right) \log\left(\frac{k}{k-j}\frac{F^{j-1}(x) - F^{k-1}(x)}{F^{j-1}(x)}\right) \mathrm{d}x.$$

The substitution $t = F^{-1}(\exp(x))$ entails that the above integral equals

$$\int_{-\infty}^{0} \frac{jk}{k-j}\left(\mathrm{e}^{jx} - \mathrm{e}^{kx}\right) \log\left(\frac{jk}{k-j}\left(\mathrm{e}^{jx} - \mathrm{e}^{kx}\right)\frac{\mathrm{e}^{-jx}}{j}\right) \mathrm{d}x = D_{KL}(f_{j,k}\|f_j).$$

Equally, one shows that

$$D_{KL}(g_j\|g_{j,k}) = D_{KL}(f_j\|f_{j,k}).$$

Therefore, the assertion is a consequence of Proposition 2.2.5. □

By repeating the arguments in the proof of Lemma 2.2.2, one easily derives the joint distribution of an arbitrary number of records. This is established by the following result.

Lemma 2.2.7. *We have for integers* $1 \le j_1 \cdots < j_d$, $d \in \mathbb{N}$, *with* $j_0 = 0$, *and* $x_1, \dots, x_d \le 0$,

$$\Pr(\eta_{j_m} \le x_m, 1 \le m \le d \mid \eta_{j_1}, \dots, \eta_{j_d} \text{ are records}) = \frac{\prod_{m=2}^{d} j_m}{\prod_{m=2}^{d}(j_m - j_{m-1})}$$
$$\times \Pr\left(\frac{\eta_m}{j_m - j_{m-1}} \le x_m, \frac{(j_{m+1} - j_m)\eta_m}{j_m - j_{m-1}} < \eta_{m+1}, 1 \le m \le d-1, \frac{\eta_d}{j_d - j_{d-1}} \le x_d\right).$$

The case of an arbitrary sequence of iid rvs $X_1, X_2, \dots$ with common continuous df F can immediately be deduced from the preceding result via the representation $X_i = F^{-1}(\exp(\eta_i))$, $i \in \mathbb{N}$.

2.3 Asymptotic Joint Distribution of Records

Let $X_1, X_2, \ldots$ be iid rvs with common df F, which is in the domain of attraction of an extreme value distribution G. From Lemma 2.2.1 we immediately obtain the following result.

Lemma 2.3.1. *Under the preceding conditions we obtain*

$$\Pr\left(\frac{X_n - b_n}{a_n} \le x \mid X_n \text{ is a record}\right) \xrightarrow[n\to\infty]{} G(x), \qquad x \in \mathbb{R}.$$

In what follows we investigate the joint asymptotic distribution of two records. We start with a sequence $\eta_1, \eta_2 \ldots$ of iid rvs that follow the standard negative exponential df $F(x) = \exp(x)$, $x \le 0$. From Lemma 2.2.2 we obtain for $x_1, x_2 \in \mathbb{R}$,

$$\begin{aligned}&\Pr\left(\eta_j \le \frac{x_1}{j}, \eta_k \le \frac{x_2}{k} \mid \eta_j \text{ and } \eta_k \text{ are records}\right)\\ &= \frac{k}{k-j}\Pr\left(\eta_1 \le x_1, \eta_2 \le \frac{k-j}{k}x_2, \frac{k-j}{j}\eta_1 < j\eta_2\right).\end{aligned}$$

We let $j = j(n)$ and $k = k(n)$ both depend on $n \in \mathbb{N}$ with

$$\lim_{n\to\infty}\frac{j}{n} = \lambda_1 > 0, \quad \lim_{n\to\infty}\frac{k}{n} = \lambda_2 > \lambda_1. \tag{2.1}$$

The next result is a consequence of Lemma 2.2.2 and elementary computations.

Proposition 2.3.2. *Under condition (2.1) we obtain*

$$\lim_{n\to\infty}\Pr\left(\eta_j \le \frac{x_1}{n}, \eta_k \le \frac{x_2}{n} \mid \eta_j \text{ and } \eta_k \text{ are records}\right) = H_{\lambda_1,\lambda_2}(x_1, x_2),$$

where

$$H_{\lambda_1,\lambda_2}(x_1,x_2) = \begin{cases} e^{\lambda_1 x_1}\left(\beta_2\, e^{(\lambda_2-\lambda_1)x_2} - \beta_1\, e^{(\lambda_2-\lambda_1)x_1}\right), & \text{if } x_1 < x_2,\\ e^{\lambda_2 x_2}, & \text{if } x_1 \ge x_2,\end{cases} \tag{2.2}$$

for all $x_1, x_2 \le 0$, $\beta_j = \lambda_j/(\lambda_2 - \lambda_1)$, $j = 1, 2$.

Remark 2.3.3. The marginal df of H_{λ_1,λ_2} are

$$H_1(x) = H_{\lambda_1,\lambda_2}(x, 0) = \beta_2\, e^{\lambda_1 x} - \beta_1\, e^{\lambda_2 x}, \quad x \le 0, \tag{2.3}$$

and

$$H_2(x) = H_{\lambda_1,\lambda_2}(0,x) = \mathrm{e}^{\lambda_2 x}, \quad x \le 0. \tag{2.4}$$

Clearly, the fact that H_2 is independent of λ_1 reflects the fact that the distribution of η_k, being a record, is not affected by the additional knowledge that η_j is a record as well, as shown in the previous section.

While H_2 is a univariate extreme value distribution, H_1 is not. Therefore, the bivariate df H_{λ_1,λ_2} is not a multivariate extreme value distribution.

In the next result we provide the marginal means, variances and the covariance of the margins of H_{λ_1,λ_2}.

Proposition 2.3.4. *Let (X,Y) be a bivariate rv with df given in (2.2). Then, we have*

$$\mathrm{E}(X) = -\lambda_1^{-1} - \lambda_2^{-1}, \quad \mathrm{E}(Y) = -\lambda_2^{-1},$$
$$\mathrm{Var}(X) = \lambda_1^{-2} + \lambda_2^{-2}, \quad \mathrm{Var}(Y) = \lambda_2^{-2}$$

and therefore

$$\begin{aligned}
\mathrm{Cov}(X,Y) &= \lambda_2^{-2}, \\
\mathrm{Cor}(X,Y) &= \frac{\lambda_1}{\sqrt{\lambda_1^2+\lambda_2^2}}, \\
\mathrm{E}\left[(X-Y)^2\right] &= \lambda_1^{-2}
\end{aligned}$$

for all $\lambda_2 > \lambda_1 > 0$.

Proof. Assume that the probability law of the pairs of the rvs (X,Y) is given by (2.2). Then, the covariance between X and Y can be computed by

$$\mathrm{Cov}(X,Y) = \int_{-\infty}^{0}\int_{-\infty}^{0} H_{\lambda_1,\lambda_2}(x,y) - H_{\lambda_1,\lambda_2}(x,0)H_{\lambda_1,\lambda_2}(0,y)\mathrm{d}x\mathrm{d}y, \tag{2.5}$$

where $H_{\lambda_1,\lambda_2}(x,y)$, $H_{\lambda_1,\lambda_2}(x,0)$ and $H_{\lambda_1,\lambda_2}(0,y)$ are given in (2.2), (2.3) and (2.4), respectively. The first term on the right-hand side of (2.5) is equal to

$$\int_{-\infty}^{0}\left(\int_{-\infty}^{y} \frac{\lambda_2\,\mathrm{e}^{(\lambda_2-\lambda_1)y+\lambda_1 x} - \lambda_1\,\mathrm{e}^{\lambda_2 x}}{\lambda_2-\lambda_1}\mathrm{d}x + \int_{y}^{0}\mathrm{e}^{\lambda_2 y}\,\mathrm{d}x\right)\mathrm{d}y$$

$$= \int_{-\infty}^{0} \frac{\lambda_2 \, e^{\lambda_2 y}}{\lambda_1(\lambda_2-\lambda_1)} dy - \int_{-\infty}^{0} \frac{\lambda_1 \, e^{\lambda_2 y}}{\lambda_2(\lambda_2-\lambda_1)} dy - \int_{-\infty}^{0} y \, e^{\lambda_2 y} \, dy$$
$$= \frac{\lambda_2^2+\lambda_1\lambda_2-2\lambda_1^2}{\lambda_1\lambda_2^2(\lambda_2-\lambda_2)}.$$

The second term on the right-hand side of (2.5) is equal to

$$\int_{-\infty}^{0} \frac{\lambda_2 \, e^{\lambda_1 x} - \lambda_1 \, e^{\lambda_2 x}}{\lambda_2-\lambda_1} dx \int_{-\infty}^{0} e^{\lambda_2 y} \, dy = \frac{\lambda_1+\lambda_2}{\lambda_1\lambda_2} \cdot \frac{1}{\lambda_2}.$$

Therefore, computing the difference between these two terms we obtain $\mathrm{Cov}(X,Y) = 1/\lambda_2^2$.

The variance of X is

$$\mathrm{Var}(X) = \mathrm{E}\left(X^2\right) - \mathrm{E}^2(X) = \frac{\lambda_1^2+\lambda_2^2}{\lambda_1^2\lambda_2^2},$$

where in particular

$$\mathrm{E}(X^2) = 2\int_{-\infty}^{0} -x \, \frac{\lambda_2 \, e^{\lambda_1 x} - \lambda_1 \, e^{\lambda_2 x}}{\lambda_2-\lambda_1} dx = \frac{2(\lambda_2^2+\lambda_1\lambda_2+\lambda_1^2)}{\lambda_1^2\lambda_2^2}$$

and

$$\mathrm{E}^2(X) = \left(-\int_{-\infty}^{0} \frac{\lambda_2 \, e^{\lambda_1 x} - \lambda_1 \, e^{\lambda_2 x}}{\lambda_2-\lambda_1} dx\right)^2 = \frac{(\lambda_1+\lambda_2)^2}{\lambda_1^2\lambda_2^2}.$$

The marginal distribution of Y is $\exp(\lambda_2 y)$, $y \le 0$, therefore its mean and variance are $1/\lambda_2$ and $1/\lambda_2^2$, respectively. Finally, combining these results we obtain the correlation

$$\mathrm{Cor}(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\,\mathrm{Var}(Y)}} = \frac{1/\lambda_2^2}{\sqrt{(\lambda_1^2+\lambda_2^2)/\lambda_1^2\lambda_2^2 \cdot 1/\lambda_2^2}} = \frac{\lambda_1}{\sqrt{\lambda_1^2+\lambda_2^2}}$$

and this completes the proof. □

The next result extends Proposition 2.3.2 to a sequence $X_1, X_2, \dots$ of iid rvs, whose joint continuous df F satisfies $F \in \mathcal{D}(G)$.

Corollary 2.3.5. *Let $X_1, X_2, \dots$ be iid copies of a rv X with a continuous distribution F. Assume that $F \in \mathcal{D}(G)$ with norming constants $a_n > 0$ and $b_n \in \mathbb{R}$, $n \in \mathbb{N}$. Then, under Condition (2.1), we have for $y_1, y_2 \in \mathbb{R}$*

$$\lim_{n\to\infty} \Pr\left(\frac{X_j-b_n}{a_n} \le y_1, \frac{X_k-b_n}{a_n} \le y_2 \mid X_j \text{ and } X_k \text{ are records}\right)$$
$$= G_{\lambda_1,\lambda_2}(y_1,y_2),$$

where

$$G_{\lambda_1,\lambda_2}(y_1,y_2) = \begin{cases} G^{\lambda_1}(y_1)\left(\beta_2 G(y_2)^{\lambda_2-\lambda_1} - \beta_1 G(y_1)^{\lambda_2-\lambda_1}\right), & \text{if } y_1 < y_2, \\ G^{\lambda_2}(y_2), & \text{if } y_1 \geq y_2. \end{cases}$$

The marginal distributions are given by

$$G_{\lambda_1,\lambda_2}(y_1,\infty) = \beta_2 G(y_1)^{\lambda_1} - \beta_1 G(y_1)^{\lambda_2}, \quad G_{\lambda_1,\lambda_2}(\infty,y_2) = G(y_2)^{\lambda_2}, \quad y \in \mathbb{R};$$

note that the second marginal is independent of λ_1.

Note that results on the limiting distribution of joint records with *known* orders in the sequence of records have been recently derived by Barakat and Elgawad (2017).

Proof. Put $\eta_m := \log(F(X_m))$, $m \in \mathbb{N}$. Then $\eta_1, \eta_2, \ldots$ are iid rvs, which follow the standard negative exponential distribution. From the fact that the functions $\log(\cdot)$ and $F(\cdot)$ are monotone we obtain

$$\begin{aligned} &\Pr\left(X_j \leq a_n y_1 + b_n, X_k \leq a_n y_2 + b_n \mid X_j \text{ and } X_k \text{ are records}\right) \\ &= \Pr\left(\eta_j \leq \log(F(a_n y_1 + b_n)), \eta_k \leq \log(F(a_n y_2 + b_n)) \mid \eta_j \text{ and } \eta_k \text{ are records}\right) \\ &= \Pr\left(\eta_j \leq \frac{n\log(F(a_n y_1 + b_n))}{n},\right. \\ &\qquad \left. \eta_k \leq \frac{n\log(F(a_n y_2 + b_n))}{n} \mid \eta_j \text{ and } \eta_k \text{ are records}\right). \end{aligned}$$

The condition

$$F^n(a_n x + b_n) \xrightarrow[n\to\infty]{} G(x), \qquad x \in \mathbb{R},$$

is equivalent to

$$n\log(F(a_n x + b_n)) \xrightarrow[n\to\infty]{} \log(G(x)), \qquad 0 < G(x) \leq 1.$$

Proposition 2.3.2 now implies

$$\begin{aligned} &\Pr\left(\eta_j \leq \frac{n\log(F(a_n y_1 + b_n))}{n}, \eta_k \leq \frac{n\log(F(a_n y_2 + b_n))}{n}\right. \\ &\qquad \left. \mid \eta_j \text{ and } \eta_k \text{ are records}\right) \xrightarrow[n\to\infty]{} H_{\lambda_1,\lambda_2}(\log(G(y_1)), \log(G(y_2))) \end{aligned}$$

which is the assertion. □

We have established the fact that the distribution of X_k, being a record, is not affected if we know in addition that X_j is a record as well. But what happens if we know, for example, that X_j, being a record, has already exceeded a fixed threshold? The answer is a straightforward consequence of our preceding results. We obtain for $y > u \in \mathbb{R}$, under the conditions of Corollary 2.3.5,

$$\begin{aligned}
&\Pr\left(\frac{X_k - b_n}{a_n} \le y \mid \frac{X_j - b_n}{a_n} > u, X_j \text{ and } X_k \text{ are records}\right) \\
&= \frac{\Pr\left(\frac{X_j - b_n}{a_n} > u, \frac{X_k - b_n}{a_n} \le y \mid X_j \text{ and } X_k \text{ are records}\right)}{\Pr\left(\frac{X_j - b_n}{a_n} > u \mid X_j \text{ and } X_k \text{ are records}\right)} \\
&\xrightarrow[n\to\infty]{} \frac{G(y)^{\lambda_2} - G^{\lambda_1}(u)\left(\beta_2 G(y)^{\lambda_2-\lambda_1} - \beta_1 G(u)^{\lambda_2-\lambda_1}\right)}{\beta_2\left(1 - G(u)^{\lambda_1}\right) - \beta_1\left(1 - G(u)^{\lambda_2}\right)}.
\end{aligned}$$

The results obtained so far can be extended to the case of an arbitrary number of records. However, computations become really hard. We report the case of the asymptotic joint df of three records.

Proposition 2.3.6. *Let $X_1, X_2, \ldots$ be iid copies of a rv X with a continuous distribution F. Assume that $F \in \mathcal{D}(G)$ with norming constants $a_n > 0$ and $b_n \in \mathbb{R}$, $n \in \mathbb{N}$. Assume also that $j = j(n)$, $k = k(n)$ and $r = r(n)$ all depending on $n \in \mathbb{N}$ with $j < k < r$ and*

$$\lim_{n\to\infty} \frac{j}{n} = \lambda_1 > 0, \quad \lim_{n\to\infty} \frac{k}{n} = \lambda_2 > \lambda_1, \quad \lim_{n\to\infty} \frac{r}{n} = \lambda_3 > \lambda_2.$$

Then, for all $\boldsymbol{y} \in \mathbb{R}^3$, we have

$$\begin{aligned}
&\Pr\left(\frac{X_j - b_n}{a_n} \le y_1, \frac{X_k - b_n}{a_n} \le y_2, \frac{X_r - b_n}{a_n} \le y_3 \mid X_j,\ X_k \text{ and } X_r \text{ are records}\right) \\
&\to \boldsymbol{G}_{\boldsymbol{\lambda}}(\boldsymbol{y}),
\end{aligned}$$

as $n \to \infty$*, where*

$$\boldsymbol{G}_{\boldsymbol{\lambda}}(\boldsymbol{y}) = \begin{cases} G(y_1)^{\lambda_1} G(y_2)^{\lambda_2-\lambda_1} \left(\beta_2\beta_6 G(y_3)^{\lambda_3-\lambda_2} - \beta_4\beta_5 G(y_2)^{\lambda_3-\lambda_1}\right) \\ -G(y_1)^{\lambda_2} \left(\beta_1\beta_6 G(y_e)^{\lambda_3-\lambda_2} - \beta_3\beta_4 G(y_1)^{\lambda_3-\lambda_2}\right) \mathrm{e}^{\lambda_2 x_1} \\ \quad \textit{if} \quad y_1 \le y_2 \le y_3 \\ \\ G(y_2)^{\lambda_2} \left(\beta_6 G(y_3)^{\lambda_3-\lambda_2} - \beta_4 G(y_2)^{\lambda_3-\lambda_2}\right), \\ \quad \textit{if} \quad y_2 \le y_1 \le y_3 \ \textit{or} \ y_2 \le y_3 \le y_1, \\ \\ \beta_2\beta_5 G(y_1)^{\lambda_1} G(y_3)^{\lambda_3-\lambda_1} - \beta_1\beta_6 G(y_1)^{\lambda_2} G(y_3)^{\lambda_3-\lambda_2} + \beta_3\beta_4 G(y_1)^{\lambda_3}, \\ \quad \textit{if} \quad y_1 \le y_3 \le y_2, \\ \\ G(y_3)^{\lambda_3}, \quad \textit{if} \quad y_3 \le y_2 \le y_1 \ \textit{or} \ y_3 \le y_1 \le y_2, \end{cases}$$

and where β_1, β_2 *are as in Proposition 2.3.2,* $\beta_3 = \lambda_1/(\lambda_3 - \lambda_1)$*,* $\beta_4 = \lambda_2/(\lambda_3 - \lambda_2)$*,* $\beta_5 = \lambda_3/(\lambda_3 - \lambda_1)$*,* $\beta_6 = \lambda_3/(\lambda_3 - \lambda_1)$*, and* $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \lambda_3)$*. Furthermore, let* $\boldsymbol{Y} = (Y_1, Y_2, Y_3)$ *be a trivariate random vector with df* $\boldsymbol{G}_{\boldsymbol{\lambda}}(\boldsymbol{y})$*. Then, the variance-covariance matrix of* $\boldsymbol{Y}$ *is*

$$\begin{pmatrix} \lambda_1^{-2} + \lambda_2^{-2} + \lambda_3^{-2} & \lambda_2^{-2} + \lambda_3^{-2} & \lambda_3^{-2} \\ & \lambda_2^{-2} + \lambda_3^{-2} & \lambda_3^{-2} \\ & & \lambda_3^{-2} \end{pmatrix}.$$

Proof. Let $\eta_1, \eta_2, \ldots$ be iid rvs with a common negative exponential distribution.

Firstly, we compute the following non-asymptotic distribution when $x_1 < x_2 < x_3$

$$
\begin{aligned}
&\Pr\left(\eta_j \le x_1, \eta_k \le x_2, \eta_r \le x_3 \mid \eta_j,\, \eta_k \text{ and } \eta_r \text{ are records}\right) \\
&= \frac{kr}{(k-j)(r-k)} \Pr\left(\eta_1 \le jx_1, \eta_2 \le (k-j)x_2, \eta_3 \le (r-k)x_3,\right. \\
&\qquad\qquad\qquad\qquad \left.(k-j)\eta_1 < j\eta_2, (r-k)\eta_2 < (k-j)\eta_3\right) \\
&= \frac{kr}{(k-j)(r-k)} \int_{-\infty}^{jx_1} \int_{\frac{k-j}{j}z_1}^{(k-j)x_2} \Pr\left(\frac{r-k}{k-j} z_2 < \eta_3 \le x_3\right) \mathrm{e}^{z_2+z_1}\, \mathrm{d}z_2 \mathrm{d}z_1 \\
&= \frac{kr}{(k-j)(r-k)} \int_{-\infty}^{jx_1} \mathrm{e}^{z_1} \left(\mathrm{e}^{(r-k)x_3}\left(\mathrm{e}^{(k-j)x_2} - \mathrm{e}^{\frac{k-j}{j}z_1}\right)\right. \\
&\qquad\qquad \left. - \frac{k-j}{r-j}\left(\mathrm{e}^{(r-j)x_2} - \mathrm{e}^{\frac{r-j}{j}z_1}\right)\right) \mathrm{d}z_1 \\
&= \frac{kr}{(k-j)(r-k)} \left(\mathrm{e}^{jx_1}\left(\mathrm{e}^{(r-k)x_3}\,\mathrm{e}^{(k-j)x_2} - \frac{k-j}{r-j}\,\mathrm{e}^{(r-j)x_2}\right) - \right. \\
&\qquad\qquad \left. \frac{j}{k}\,\mathrm{e}^{kx_1}\,\mathrm{e}^{(r-k)x_3} + \frac{j(k-j)}{r(r-j)}\,\mathrm{e}^{rx_1}\right).
\end{aligned}
$$

The cases of $x_2 < x_1 < x_3$ is obtained from the expression of the above formula by substituting x_2 in x_1. Similarly the case $x_1 < x_3 < x_2$ is obtained by substituting x_3 in x_2 and lastly the case $x_3 < x_2 < x_1$ is obtained by substituting x_3 in both x_1 and x_2. Then, the asymptotic distribution is easily obtained by computing

$$
\lim_{n\to\infty} \Pr\left(\eta_j \le x_1/n, \eta_k \le x_2/n, \eta_r \le x_3/n \mid \eta_j,\, \eta_k \text{ and } \eta_r \text{ are records}\right)
$$

The case of an arbitrary distribution can be deduced by following the same reasoning of Corollary 2.3.5 and therefore the first assertion is derived.

We compute the variance-covariance matrix. Note that the bivariate and univariate marginal distribution functions of (Y_1, Y_2) are

$$
F_{\lambda_1,\lambda_2}(x_1,x_2) = \begin{cases}
\left(\frac{\lambda_2\lambda_3}{(\lambda_2-\lambda_1)(\lambda_3-\lambda_2)}\,\mathrm{e}^{(\lambda_2-\lambda_1)x_2} - \frac{\lambda_2\lambda_3}{(\lambda_3-\lambda_2)(\lambda_3-\lambda_1)}\,\mathrm{e}^{(\lambda_3-\lambda_1)x_2}\right)\mathrm{e}^{\lambda_1 x_1} \\
-\frac{\lambda_1\lambda_3}{(\lambda_2-\lambda_1)(\lambda_3-\lambda_2)}\,\mathrm{e}^{\lambda_2 x_1} - \frac{\lambda_1\lambda_2}{(\lambda_3-\lambda_1)(\lambda_3-\lambda_2)}\,\mathrm{e}^{\lambda_3 x_1}, \quad \text{if} \quad x_1 \le x_2, \\
\frac{\lambda_3\,\mathrm{e}^{\lambda_2 x_2} - \lambda_2\,\mathrm{e}^{\lambda_3 x_2}}{\lambda_3-\lambda_2}, \quad \text{if} \quad x_2 \le x_1,
\end{cases}
$$

$$F_{\lambda_1}(x_1) = \frac{\lambda_2\lambda_3}{(\lambda_2-\lambda_1)(\lambda_3-\lambda_1)}\,\mathrm{e}^{\lambda_1 x_1} - \frac{\lambda_1\lambda_3}{(\lambda_2-\lambda_1)(\lambda_3-\lambda_2)}\,\mathrm{e}^{\lambda_2 x_1}$$
$$- \frac{\lambda_1\lambda_2}{(\lambda_3-\lambda_1)(\lambda_3-\lambda_2)}\,\mathrm{e}^{\lambda_3 x_1},\ x_1 \le 0,$$
$$F_{\lambda_2}(x_2) = \frac{\lambda_3\,\mathrm{e}^{\lambda_2 x_2} - \lambda_2\,\mathrm{e}^{\lambda_3 x_2}}{\lambda_3-\lambda_2},\ x_2 \le 0.$$

Hoeffding's covariance identity implies that

$$\begin{aligned}\mathrm{Cov}(Y_1,Y_2) &= \int_{(-\infty,0]^2} F_{\lambda_1,\lambda_2}(x_1,x_2)\mathrm{d}x_1\mathrm{d}x_2 - \int_{-\infty}^{0} F_{\lambda_1}(x_1)\mathrm{d}x_1 \int_{-\infty}^{0} F_{\lambda_2}(x_2)\mathrm{d}x_2\\ &= \frac{\lambda_3}{\lambda_1(\lambda_2-\lambda_1)(\lambda_3-\lambda_2)} - \frac{\lambda_2}{\lambda_1(\lambda_3-\lambda_1)(\lambda_3-\lambda_2)}\\ &\quad - \frac{\lambda_1\lambda_3}{\lambda_2^2(\lambda_2-\lambda_1)(\lambda_3-\lambda_2)} + \frac{\lambda_1\lambda_2}{\lambda_3^2(\lambda_3-\lambda_1)(\lambda_3-\lambda_2)}\\ &\quad + \frac{\lambda_3}{\lambda_2^2(\lambda_3-\lambda_2)} - \frac{\lambda_2}{\lambda_3^2(\lambda_3-\lambda_2)}\\ &\quad - \left(\frac{1}{\lambda_1}+\frac{1}{\lambda_2}+\frac{1}{\lambda_3}\right)\left(\frac{1}{\lambda_2}+\frac{1}{\lambda_3}\right).\end{aligned}$$

By straightforward simplifications we obtain $\mathrm{Cov}(Y_1,Y_2) = \lambda_2^{-2} + \lambda_3^{-2}$. The other covariances are computed in a similar way. The expected values and the variances are derived by elementary computations. Therefore, the second assertion is also proven. □

Under Condition (2.1), another application of Lemma 2.2.7 yields the following results.

Theorem 2.3.7. *Let $\eta_1, \eta_2, \dots$ be independent and standard negative exponential distributed rvs. Assume that $j_i = j_i(n) \in \mathbb{N}$, $i = 1, 2, \dots$, $n = 2, 3, \dots$ are sequences of integers satisfying*

$$\lim_{n\to\infty} \frac{j_i}{n} = \lambda_i > 0,$$

with $0 < \lambda_1 < \lambda_2 < \cdots$. Then, under these conditions we have

$$\lim_{n\to\infty} \Pr\left(\eta_k - \eta_j \le y/n \mid \eta_j \text{ and } \eta_k \text{ are records}\right) = 1 - \mathrm{e}^{-\lambda_1 y},$$

$$\lim_{n\to\infty} \Pr\left(\eta_{j_{i+1}} - \eta_{j_i} \le y_i/n, 1 \le i \le s \mid \eta_{j_1} \dots \eta_{j_s} \text{ are records}\right) = \prod_{i=1}^{s}\left(1 - \mathrm{e}^{-\lambda_i y_i}\right),$$

$$\lim_{n\to\infty} \Pr\left(\eta_j \le x/n, \eta_k - \eta_j \le y/n \mid \eta_j \text{ and } \eta_k \text{ are records}\right) = Q_{\lambda_1,\lambda_2}(x,y),$$

where

$$Q_{\lambda_1,\lambda_2}(x,y) = \begin{cases} \beta_1 \left(\mathrm{e}^{(\lambda_2-\lambda_1)y} - 1\right)\mathrm{e}^{\lambda_2 x}, & \text{if } |x| \geq y, \\ \beta_2\,\mathrm{e}^{\lambda_1 x} - \beta_1\,\mathrm{e}^{\lambda_2 x} - \mathrm{e}^{-\lambda_1 y}, & \text{if } |x| < y, \end{cases} \tag{2.6}$$

for all $x \leq 0$, $y, y_1, \dots, y_s > 0$ *and* $m \in \mathbb{N}$.

REMARK 2.3.8. *The marginal distributions functions of* (2.6) *are*

$$Q_{\lambda_1,\lambda_2}(x) = Q_{\lambda_1,\lambda_2}(x,\infty) = \beta_2\,\mathrm{e}^{\lambda_1 x} - \beta_1\,\mathrm{e}^{\lambda_2 x}, \quad x \leq 0,$$

and

$$Q_{\lambda_1}(y) = Q_{\lambda_1,\lambda_2}(0,y) = 1 - \mathrm{e}^{-\lambda_1 y}, \quad y > 0.$$

These results mean that the increments $Y_1, \dots, Y_n$ among records are independent but not identically distributed. Furthermore, a generic record η_j, $j = 1, 2, \dots$ and the increment between two records η_j and η_k, $k > j$ are not independent.

Proof. For all $y_1, \dots, y_s > 0$, by following the same reasoning in the proof of Lemma 2.2.2 we have

$$\begin{aligned}
&\Pr\left(\eta_{j_2} - \eta_{j_1} > y_1, \dots, \eta_{j_s} - \eta_{j_{s-1}} > y_{s-1} \mid \eta_{j_1} \dots \eta_{j_s} \text{ are records}\right) \\
&= \frac{\prod_{m=2}^{s} j_m}{\prod_{m=2}^{s}(j_m - j_{m-1})} \Pr\left(\frac{\eta_2}{j_2 - j_1} - \frac{\eta_1}{j_1} > z_1, \dots, \frac{\eta_s}{j_s - j_{s-1}} - \frac{\eta_{s-1}}{j_{s-1}} > z_{s-1}\right) \\
&= \Pr\left(\eta_{j_2} - \eta_{j_1} > y_1, \dots, \eta_{j_s} - \eta_{j_{s-1}} > y_{s-1} \mid \eta_{j_1} \dots \eta_{j_s} \text{ are records}\right) \\
&= \frac{\prod_{m=2}^{s} j_m}{\prod_{m=2}^{s}(j_m - j_{m-1})} \cdot A
\end{aligned}$$

where

$$A = \int_{-\infty}^{0} \int_{-\infty}^{(j_{s-1}-j_{s-2})\left(\frac{z_s}{j_s - j_{s-1}} - y_{s-1}\right)} \cdots \int_{-\infty}^{(j_2-j_1)\left(\frac{z_3}{j_3-j_2} - y_2\right)} \Pr\left(\eta_1 < j_1\left(\frac{z_2}{j_2 - j_1} - y_1\right)\right) \prod_{i=2}^{s} \mathrm{e}^{z_i}\,\mathrm{d}z_i.$$

We show by induction that

$$\int_{-\infty}^{(j_m-j_{m-1})\left(\frac{z_{m+1}}{j_{m+1}-j_m} - y_m\right)} \cdots \int_{-\infty}^{(j_2-j_1)\left(\frac{z_3}{j_3-j_2} - y_2\right)} \mathrm{e}^{\frac{j_2}{j_2-j_1} z_2}\,\mathrm{e}^{-j_1 y_1} \prod_{i=2}^{m} \mathrm{e}^{z_i}\,\mathrm{d}z_i$$

$$= \frac{\prod_{i=2}^{m}(j_i - j_{i-1})}{\prod_{i=2}^{m} j_i} \prod_{i=1}^{m} \mathrm{e}^{-j_i y_i}\, \mathrm{e}^{\frac{j_m z_{m+1}}{j_{m+1}-j_m}} .$$

At the step 1 we have

$$\int_{-\infty}^{(j_2-j_1)\left(\frac{z_3}{j_3-j_2}-y_2\right)} \mathrm{e}^{\frac{j_2}{j_2-j_1}z_2}\, \mathrm{e}^{-j_1 y_1}\, \mathrm{d}z_2 = \frac{j_2-j_1}{j_2}\, \mathrm{e}^{-j_1 y_1 - j_2 y_2}\, \mathrm{e}^{\frac{j_2 z_3}{j_3-j_2}} .$$

True for m. At the step $m+1$ we have

$$\int_{-\infty}^{(j_{m+1}-j_m)\left(\frac{z_{m+2}}{j_{m+2}-j_{m+1}}-y_{m+1}\right)} \cdots \int_{-\infty}^{(j_2-j_1)\left(\frac{z_3}{j_3-j_2}-y_2\right)} \mathrm{e}^{\frac{j_2}{j_2-j_1}z_2}\, \mathrm{e}^{-j_1 y_1} \prod_{i=2}^{m+1} \mathrm{e}^{z_i}\, \mathrm{d}z_i$$

$$= \int_{-\infty}^{(j_{m+1}-j_m)\left(\frac{z_{m+2}}{j_{m+2}-j_{m+1}}-y_{m+1}\right)} \frac{\prod_{i=2}^{m}(j_i - j_{i-1})}{\prod_{i=2}^{m} j_i} \prod_{i=1}^{m} \mathrm{e}^{-j_i y_i}\, \mathrm{e}^{\frac{j_m z_{m+1}}{j_{m+1}-j_m}}\, \mathrm{e}^{z_{m+1}}\, \mathrm{d}z_{m+1}$$

$$= \frac{\prod_{i=2}^{m+1}(j_i - j_{i-1})}{\prod_{i=2}^{m+1} j_i} \prod_{i=1}^{m+1} \mathrm{e}^{-j_i y_i}\, \mathrm{e}^{\frac{j_{m+1} z_{m+2}}{j_{m+2}-j_{m+1}}} .$$

As a consequence

$$A = \frac{\prod_{i=2}^{s}(j_i - j_{i-1})}{\prod_{i=2}^{s} j_i} \prod_{i=1}^{s-1} \mathrm{e}^{-j_i y_i} .$$

We obtain

$$\Pr\left(\eta_{j_2} - \eta_{j_1} > y_1, \ldots, \eta_{j_s} - \eta_{j_{s-1}} > y_{s-1} \mid \eta_{j_1} \ldots \eta_{j_s} \text{ are records}\right) = \prod_{i=1}^{s-1} \mathrm{e}^{-j_i y_i},$$

which shows that the increments among records are independent exponentials but with different parameters $0 < \lambda_1 < \lambda_2 < \cdots$. By the assumptions we have that $(1 - \mathrm{e}^{j_i y_i/n}) \to (1 - \mathrm{e}^{\lambda_i y_i})$ as $n \to \infty$ for any $i = 1, \ldots, s$ and therefore the second result is proven.

Next, for $x \leq 0$, $y \geq 0$ we have

$$\begin{aligned} Q(x,y) &:= \Pr\left(\eta_j \leq x, \eta_k - \eta_j \leq y \mid \eta_j \text{ and } \eta_k \text{ are records}\right) \\ &= jk \Pr\left(\eta_j \leq x, \eta_k - \eta_j \leq y, \eta_j > \frac{\eta_1}{j-1}, \eta_k > \max\left(\eta_j, \frac{\eta_2}{k-j-1}\right)\right) \end{aligned}$$

When $x \leq -y$,

$$\begin{aligned} Q(x,y) &= jk \int_{-\infty}^{x} \int_{z_j}^{z_j+y} \mathrm{e}^{(k-j)z_k}\, \mathrm{e}^{j z_j}\, \mathrm{d}z_k \mathrm{d}z_j \\ &= \frac{jk}{k-j}(\mathrm{e}^{(k-j)y} - 1) \int_{-\infty}^{x} \mathrm{e}^{k z_j}\, \mathrm{d}z_j = \frac{j}{k-j}(\mathrm{e}^{(k-j)y} - 1)\, \mathrm{e}^{kx} . \end{aligned}$$

Therefore, $Q(x/n, y/n) \to \beta_1 \left(\mathrm{e}^{(\lambda_2-\lambda_1)y} - 1\right) \mathrm{e}^{\lambda_2 x}$ as $n \to \infty$. When $x > -y$

$$\begin{aligned}
Q(x,y) &= jk\left(\int_{-\infty}^{-y}\int_{z_j}^{z_j+y} \mathrm{e}^{(k-j)z_k}\,\mathrm{e}^{jz_j}\,\mathrm{d}z_k\mathrm{d}z_j + \int_{-y}^{x}\int_{z_j}^{0} \mathrm{e}^{(k-j)z_k}\,\mathrm{e}^{jz_j}\,\mathrm{d}z_k\mathrm{d}z_j\right) \\
&= \frac{j}{k-j}(\mathrm{e}^{(k-j)y} - 1)\,\mathrm{e}^{-ky} + \frac{jk}{k-j}\left(\frac{\mathrm{e}^{jx} - \mathrm{e}^{-jy}}{j} - \frac{\mathrm{e}^{kx} - \mathrm{e}^{-ky}}{k}\right) \\
&= \frac{k}{k-j}\,\mathrm{e}^{jx} - \frac{j}{k-j}\,\mathrm{e}^{kx} - \mathrm{e}^{-jy}\,.
\end{aligned}$$

Therefore $Q(x/n, y/n) \to \beta_2\,\mathrm{e}^{\lambda_1 x} - \beta_1\,\mathrm{e}^{\lambda_2 x} - \mathrm{e}^{-\lambda_1 y}$ as $n \to \infty$. □

Chapter 3

On Multivariate Records from Random Vectors with Independent Components

3.1 Introduction

Let $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots$ be independent copies of a random vector $\boldsymbol{X} \in \mathbb{R}^d$ with distribution function F. We assume that the margins $F_i, 1 \leq i \leq d$, of F are continuous univariate distribution functions. This is equivalent of assuming the condition that F itself is a continuous distribution function.

In this chapter we are interested in *complete* records. The d-dimensional random vector (rv) $\boldsymbol{X}_j$ is a complete record if each of its components is a record, i.e.,

$$\boldsymbol{X}_j > \max_{1 \leq i \leq j-1} \boldsymbol{X}_i,$$

where the maximum is taken componentwise. All our operations on vectors $\boldsymbol{x} = (x_1, \ldots, x_d)$, $\boldsymbol{y} = (y_1 \ldots, y_d)$, such as $\boldsymbol{x} < \boldsymbol{y}$, are meant componentwise. Clearly, $\boldsymbol{X}_1$ is a complete record.

Multivariate records have not been discussed that extensively, yet they have been approached by Goldie and Resnick (1989), Goldie and Resnick (1995) and Arnold, Balakrishnan, and Nagaraja (1998, Chapter 8), for instance. For supplementary material on multivariate and functional records we refer to the thesis by Zott (2016) and the references cited therein.

We focus on the case where the components of $\boldsymbol{X}$ are independent. Then, clearly, various of the results for univariate vectors carry over to the multivariate case. In particular the preceding result carries over: Put

$$R_m^{\mathrm{CR}} := \mathbb{1}(\boldsymbol{X}_m \text{ is a complete record}).$$

Then, the indicator functions $R_1^{\mathrm{CR}}, R_2^{\mathrm{CR}}, \dots$ are independent with

$$\Pr(R_m^{\mathrm{CR}} = 1) = m^{-d}, \qquad m \in \mathbb{N}. \tag{3.1}$$

However, from equations (1.6) and (3.1) we immediately deduce a first difference between the theory of univariate and multivariate records. If the joint distribution function of the sequence of iid univariate random variables is continuous, then the total number of records in this series is infinite with probability one. On the other hand, by equation (3.1) we have

$$\mathrm{E}\left(\sum_{m\in\mathbb{N}} I_m\right) = \sum_{m\in\mathbb{N}} \mathrm{E}(I_m) = \sum_{m\in\mathbb{N}} m^{-d} < \infty \tag{3.2}$$

if $d \geq 2$. As a consequence, the total number of complete records $\sum_{m\in\mathbb{N}} I_m$ is finite with probability one. Hence, in case $d \geq 2$, there is a terminal complete record in the series $\boldsymbol{X}_1, \boldsymbol{X}_2, \dots$ In Section 3.2 we compute the distribution of the random total number of complete records and we investigate the distribution of the terminal record. In Section 3.3 we study the sequence of waiting times for the complete records. Such a sequence forms a Markov chain, similarly to the univariate case, but in higher dimensions the state infinity is an absorbing element of this state space.

Suppose that the components of $\boldsymbol{X}$ are not independent, but that its distribution function is in the max-domain of attraction of a multivariate extreme value distribution. Goldie and Resnick (1989, Theorem 5.3) proved in case $d = 2$ that the total number of complete records is finite if and only if the limiting extreme value distribution has independent components. Assuming that the components of $\boldsymbol{X}$ are independent, we only require continuity of its distribution function, we do not requite that it is in the max-domain of attraction of an extreme value distribution.

3.2 Terminal record

Let

$$L := \sup \left\{ m \in \mathbb{N} : R_m^{\text{CR}} = 1 \right\},$$

which is the index of the ultimate complete record in the sequence $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots$ If $d \geq 2$, then we know from equation (3.2) that $\Pr(L < \infty) = 1$. In the next Lemma we compute the distribution of L.

Lemma 3.2.1. *For $d \geq 2$ and $k \in \mathbb{N}$*

$$p_k = \Pr(L = k) = \frac{1}{k^d} \prod_{m \geq k+1} \left(1 - \frac{1}{m^d}\right). \tag{3.3}$$

In particular, when $d = 2$:

$$p_k = \frac{1}{k(k+1)}.$$

Proof. The independence of the indicator functions I_m, $m \in \mathbb{N}$, together with equation (3.1) imply the first assertion.

In the case $d = 2$ we obtain

$$p_k = \frac{1}{k^2} \lim_{N \to \infty} \prod_{m=k+1}^{N} \frac{(m-1)(m+1)}{m^2} = \frac{1}{k^2} \lim_{N \to \infty} \frac{k(N+1)}{N(k+1)} = \frac{1}{k(k+1)}.$$

□

The first observation $\boldsymbol{X}_1$ is a record by definition. By the preceding result, in dimension $d = 2$, $\boldsymbol{X}_1$ is already the terminal complete record with probability $p_1 = 1/2$. The next observation $\boldsymbol{X}_2$ is with probability $1/4$ a complete record, and it is with probability $1/6$ the terminal complete record. From equation (3.2) and Lemma 3.2.1 we have that

$$1 = \Pr(L < \infty) = \sum_{k \in \mathbb{N}} p_k = \frac{1}{k^d} \prod_{m \geq k+1} \left(1 - \frac{1}{m^d}\right)$$

which, taken as a purely mathematical formula, is a nice by-product.

The probability $p_1 = p_1(d)$ increases as the dimension increases, whereas $p_k = p_k(d)$, $k \geq 2$, decreases. This is the content of the next Lemma.

Lemma 3.2.2. *We have*

$$\lim_{d\to\infty} p_1(d) = \lim_{d\to\infty} \prod_{m\geq 2} \left(1 - \frac{1}{m^d}\right) = 1,$$

whereas, for $k \geq 2$, we have

$$\lim_{d\to\infty} p_k(d) = \lim_{d\to\infty} \frac{1}{k^d} \prod_{m\geq k+1} \left(1 - \frac{1}{m^d}\right) = 0.$$

Proof. The second assertion is immediate from the bound $\prod_{m\geq k+1}\left(1 - \frac{1}{m^d}\right) \leq 1$. The first assertion is a consequence of the equation $\sum_{k\in\mathbb{N}} p_k = 1$ and the following bound, valid for $d \geq 4$,

$$\sum_{k\geq 2} p_k = \sum_{k\geq 2} \frac{1}{k^d} \prod_{m\geq k+1} \left(1 - \frac{1}{m^d}\right) \leq \frac{1}{2^{d-2}} \sum_{k\geq 2} \frac{1}{k^2} \xrightarrow[d\to\infty]{} 0.$$

□

Lemma 3.2.1 implies that the expected arrival time for the final complete record is

$$\mathrm{E}(L) = \sum_{k\in\mathbb{N}} k p_k = \sum_{k\in\mathbb{N}} \frac{1}{k^{d-1}} \prod_{m\geq k+1} \left(1 - \frac{1}{m^d}\right).$$

Therefore, we have

$$\mathrm{E}(L) = \sum_{k\in\mathbb{N}} \frac{1}{k+1} = \infty$$

when $d = 2$, while

$$\mathrm{E}(L) \leq \sum_{k\in\mathbb{N}} \frac{1}{k^{d-1}} \leq \infty$$

for $d \geq 3$.

Finally, by repeating the arguments in the proof of Lemma 3.2.2, we have

$$\mathrm{E}(L) \xrightarrow[d\to\infty]{} 1.$$

Let $\boldsymbol{\eta}$ be a d-dimensional rv with independent components $\eta_1, \dots, \eta_d$, each following a standard negative exponential distribution, i.e., $\Pr(\eta_i \le x) = \exp(x)$, $x \le 0$, for all $i \le d$. In what follows we investigate the distribution of the terminal record, i.e., we study $\Pr(\boldsymbol{\eta}_L \le \boldsymbol{x})$, where L denotes again the random index of the terminal record and $\boldsymbol{x} = (x_1, \dots, x_d) \in (-\infty, 0]^d$. We have a closed formula for $\Pr(\boldsymbol{\eta}_L \le \boldsymbol{x})$, see Theorem 3.2.5. However, we first want to verify the following conjecture: Let $L = L(d)$ be the random index of the terminal record which depends on the dimension d. From Lemma 3.2.2 we know that $\Pr(L(d) = 1) = p_1(d) \xrightarrow[d\to\infty]{} 1$. Therefore, one would expect that

$$\Pr(\boldsymbol{\eta}_L \le \boldsymbol{x}) \approx \Pr(\boldsymbol{\eta}_1 \le \boldsymbol{x}) = \exp\left(\sum_{i=1}^d x_i\right),$$

when d gets large. This conjecture is verified in the next result. To ease the notation we drop the dependence on d, wherever it causes no ambiguities.

Proposition 3.2.3. *Let $x_1, x_2, \dots$ be a sequence of numbers in $(-\infty, 0]$.*

(i) If

$$\sum_{i=1}^\infty x_i \in (-\infty, 0],$$

then, with $\boldsymbol{x}_d := (x_1, \dots, x_d)$, $d \in \mathbb{N}$, we have

$$\Pr(\boldsymbol{\eta}_L \le \boldsymbol{x}_d) \xrightarrow[d\to\infty]{} \exp\left(\sum_{i=1}^\infty x_i\right) = \lim_{d\to\infty} \Pr(\boldsymbol{\eta}_1 \le \boldsymbol{x}_d).$$

(ii) If

$$\lim_{d\to\infty} \sum_{i=1}^d x_i = -\infty,$$

but such that

$$\limsup_{d\to\infty} \left| \frac{1}{d} \sum_{i=1}^d x_i \right| < \log 2 \tag{3.4}$$

then

$$\frac{\Pr(\boldsymbol{\eta}_L \le \boldsymbol{x}_d)}{\Pr(\boldsymbol{\eta}_1 \le \boldsymbol{x}_d)} \xrightarrow[d\to\infty]{} 1.$$

Proof. We have

$$\Pr(\boldsymbol{\eta}_L \leq \boldsymbol{x}_d) = \Pr(\boldsymbol{\eta}_L \leq \boldsymbol{x}_d, L = 1) + \Pr(\boldsymbol{\eta}_L \leq \boldsymbol{x}_d, L \geq 2)$$

with

$$\begin{aligned}
\Pr(\boldsymbol{\eta}_L \leq \boldsymbol{x}_d, L \geq 2) &= \sum_{k=2}^{\infty} \Pr(\boldsymbol{\eta}_L \leq \boldsymbol{x}_d, L = k) \\
&\leq \sum_{k=2}^{\infty} \Pr(\boldsymbol{\eta}_k \leq \boldsymbol{x}_d, \boldsymbol{\eta}_k \text{ is a complete record}) \\
&= \sum_{k=2}^{\infty} \Pr(\boldsymbol{\eta}_k \leq \boldsymbol{x}_d \mid \boldsymbol{\eta}_k \text{ is a complete record}) \frac{1}{k^d} \\
&= \sum_{k=2}^{\infty} \exp\left(k \sum_{i=1}^{d} x_i\right) \frac{1}{k^d} \\
&= \exp\left(\sum_{i=1}^{d} x_i\right) \sum_{k=2}^{\infty} \exp\left((k-1) \sum_{i=1}^{d} x_i\right) \frac{1}{k^d} \\
&= \exp\left(\sum_{i=1}^{d} x_i\right) o(1).
\end{aligned} \tag{3.5}$$

In the preceding list we used the fact that the *univariate* distribution function $\Pr(\eta_k \leq x \mid \eta_k \text{ is a record})$ equals $\exp(kx)$, $x \leq 0$, $k \in \mathbb{N}$, as established in Falk, Chokami, and Padoan (2018).

From equation (3.5) we obtain

$$\begin{aligned}
\Pr(\boldsymbol{\eta}_L \leq \boldsymbol{x}_d) &= \Pr(\boldsymbol{\eta}_L \leq \boldsymbol{x}_d, L = 1) + \exp\left(\sum_{i=1}^{d} x_i\right) o(1) \\
&= \Pr(\boldsymbol{\eta}_1 \leq \boldsymbol{x}_d) - \Pr(\boldsymbol{\eta}_1 \leq \boldsymbol{x}_d, L \geq 2) + \exp\left(\sum_{i=1}^{d} x_i\right) o(1)
\end{aligned}$$

As $\Pr(L \geq 2) \xrightarrow[d \to \infty]{} 0$, this implies the first assertion.

Next, suppose next that

$$\sum_{i=1}^{\infty} x_i = -\infty.$$

We have to show that

$$\Pr(\boldsymbol{\eta}_1 \leq \boldsymbol{x}_d, L \geq 2) = \exp\left(\sum_{i=1}^{d} x_i\right) o(1) \tag{3.6}$$

as well. Hoelder's inequality implies with $p, q \geq 1$, $p^{-1} + q^{-1} = 1$,

$$\begin{aligned}
\Pr(\boldsymbol{\eta}_L \leq \boldsymbol{x}_d, L \geq 2) &= \mathrm{E}\left(\mathbb{1}(\boldsymbol{\eta}_1 \leq \boldsymbol{x}_d)\mathbb{1}(L \geq 2)\right) \\
&\leq \Pr(\boldsymbol{\eta}_1 \leq \boldsymbol{x}_d)^{1/p}\Pr(L \geq 2)^{1/q} \\
&= \exp\left(\frac{1}{p}\sum_{i=1}^{d} x_i\right)\Pr(L \geq 2)^{1/q}
\end{aligned}$$

where

$$\begin{aligned}
\Pr(L \geq 2) &= \Pr(L = 2) + \sum_{k=3}^{\infty}\Pr(L = k) \\
&\leq 2^{-d} + \sum_{k=3}^{\infty} k^{-d} \leq 2^{-d} + \int_2^{\infty} x^{-d}\mathrm{d}x \\
&= 2^{-d} + \frac{2^{-d+1}}{d-1} \leq \frac{3}{2^d}.
\end{aligned}$$

Thus, we obtain

$$\begin{aligned}
\Pr(\boldsymbol{\eta}_1 \leq \boldsymbol{x}_d, L \geq 2) &\leq \exp\left(\frac{1}{p}\sum_{i=1}^{d} x_i\right)\frac{3^{1/q}}{2^{d/p}} \\
&= \exp\left(\sum_{i=1}^{d} x_i\right)\exp\left(\left(\frac{1}{p} - 1\right)\sum_{i=1}^{d} x_i - \frac{d}{q}\log 2\right)3^{1/q} \\
&= \exp\left(\sum_{i=1}^{d} x_i\right)\exp\left(-\frac{1}{q}\sum_{i=1}^{d} x_i - \frac{d}{q}\log 2\right)3^{1/q} \\
&= \exp\left(\sum_{i=1}^{d} x_i\right)\exp\left(\frac{d}{q}\left(-\frac{1}{d}\sum_{i=1}^{d} x_i - \log 2\right)\right)3^{1/q} \\
&= \exp\left(\sum_{i=1}^{d} x_i\right)o(1)
\end{aligned}$$

by condition (3.2). This proves the second assertion as well. □

By assuming the componentwise representation $X_i = F_i^{-1}(\exp(\eta_i))$, $i = 1, \dots, d$, $d \in \mathbb{N}$, for each component $i = 1, \dots, d$, the preceding result immediately carries over to a sequence $\boldsymbol{X}_1, \boldsymbol{X}_2, \dots$ of independent copies of a rv $\boldsymbol{X}$ with independent components and univariate continuous marginal df $F_1, \dots, F_d$.

Corollary 3.2.4. *Let $y_1, y_2, \dots$ be a sequence of number in $\mathbb{R}$.*

(i) If

$$\prod_{i=1}^{\infty} F_i(y_i) \in (0,1),$$

then

$$\Pr(\boldsymbol{X}_L \le \boldsymbol{y}_d) \underset{d\to\infty}{\longrightarrow} \prod_{i=1}^{\infty} F_i(y_i) = \lim_{d\to\infty} \Pr(\boldsymbol{X}_1 \le \boldsymbol{y}_d).$$

(ii) If

$$\prod_{i=1}^{\infty} F_i(y_i) = 0, \text{ but such that } \liminf_{d\to\infty} \left(\prod_{i=1}^{d} F_i(y_i)\right)^{1/d} > \frac{1}{2},$$

then

$$\frac{\Pr(\boldsymbol{X}_L \le \boldsymbol{y}_d)}{\Pr(\boldsymbol{X}_1 \le \boldsymbol{y}_d)} \underset{d\to\infty}{\longrightarrow} 1.$$

In the final result of this section we derive the exact distribution of the terminal complete record for fixed dimension $d \ge 2$. We suppose again a sequence $\boldsymbol{X}_1, \boldsymbol{X}_2, \dots$ of independent copies of a rv $\boldsymbol{X} \in \mathbb{R}^d$ with independent components and continuous univariate marginal df $F_1, \dots, F_d$.

Theorem 3.2.5. *The distribution function of the final complete record is*

$$\Pr(\boldsymbol{X}_L \le \boldsymbol{x}_d) = \sum_{k=1}^{\infty} \left(\frac{\prod_{i=1}^{d} u_i^k}{k^d} - \sum_{K \subseteq \mathcal{J}} (-1)^{|K|-1} \prod_{i=1}^{d} \left(\sum_{r \in K'} \frac{u_i^r}{r \prod_{s \ne r \in K'} (s-r)} + \frac{1}{\prod_{r \in K'} r} \right)\right),$$

where $u_i = F_i(x_i)$, $\mathcal{J} = \{k+1, k+2, \dots\}$, $K \subseteq \mathcal{J}$, $K' = \{k\} \cup K$ and $|K|$ is the total number of elements in the set K.

Proof of Theorem 3.2.5. Without loss of generality we provide the proof with uniform margins $U_1, U_2 \dots$ We look for the solution of

$$\Pr(\boldsymbol{U}_L \le \boldsymbol{u}_d) = \sum_{k=1}^{\infty} \Pr\left(\boldsymbol{U}_k \le \boldsymbol{u}_d \mid R_k^{\mathrm{CR}} = 1, \bigcap_{m=k+1}^{\infty} \{R_m^{\mathrm{CR}} = 0\}\right) \Pr(L = k),$$

and $\Pr(L = k) = p_k$. The probability of the conditioning event is given by (3.3), therefore we only need to compute

$$\Pr\left(\boldsymbol{U}_k \leq \boldsymbol{u}_d, R_k^{\mathrm{CR}} = 1, \bigcap_{m=k+1}^{\infty} \{R_m^{\mathrm{CR}} = 0\}\right)$$
$$= \Pr\left(\boldsymbol{U}_k \leq \boldsymbol{u}_d, R_k^{\mathrm{CR}} = 1\right) - \Pr\left(\boldsymbol{U}_k \leq \boldsymbol{u}, R_k^{\mathrm{CR}} = 1, \left(\bigcap_{m=k+1}^{\infty} \{R_m^{\mathrm{CR}} = 0\}\right)^{\mathrm{c}}\right)$$

Since the components of $\boldsymbol{U}$ are independent, it is easy to see that

$$\Pr\left(\boldsymbol{U}_k \leq \boldsymbol{u}, R_k^{\mathrm{CR}} = 1\right) = \frac{\prod_{i=1}^{d} u_i^k}{k^d}.$$

By means of the inclusion-exclusion principle, we have that

$$\Pr\left(\boldsymbol{U}_k \leq \boldsymbol{u}_d, R_k^{\mathrm{CR}} = 1, \left(\bigcap_{m=k+1}^{\infty} \{R_m^{\mathrm{CR}} = 0\}\right)^{\mathrm{c}}\right)$$
$$= \sum_{K \subseteq \mathcal{J}} (-1)^{|K|-1} \Pr\left(\boldsymbol{U}_k \leq \boldsymbol{u}_d, R_k^{\mathrm{CR}} = 1, R_t^{\mathrm{CR}} = 1, t \in K\right)$$
$$= \sum_{K \subseteq \mathcal{J}} (-1)^{|K|-1} \prod_{i=1}^{d} \Pr\left(U_{k,i} \leq u_i, I_k = 1, I_t = 1, t \in K\right)$$

where $K = \{j_1, \dots, j_{|K|}\} \subseteq \mathcal{J} = \{k+1, k+2, \dots\}$. Note that

$$\Pr\left(U_{k,i} \leq u_i, e_k = 1, e_t = 1, t \in K\right)$$
$$= \int \cdots \int_{0 \leq \min(u_i, z_1) < \cdots < z_{|K|} \leq 1} z^{k-1} \prod_{t=1}^{|K|} z_t^{j_t - j_{t-1} - 1} \,\mathrm{d}z \,\mathrm{d}z_1 \dots \mathrm{d}z_{|K|}.$$

We compute the previous probability by using the induction principle. We claim that

$$A_m = \int \cdots \int_{0 \leq \min(u_i, z_1) < \cdots < z_m} z^{k-1} \prod_{t=1}^{m-1} z_t^{j_t - j_{t-1} - 1} \,\mathrm{d}z \,\mathrm{d}z_1 \dots \mathrm{d}z_{|K|-1}$$
$$= \sum_{t=0}^{m-1} (-1)^t \frac{u_i^{j_t}}{\prod_{r=0}^{t-1}(j_t - j_r) j_t \prod_{r=t+1}^{m-1}(j_r - j_t)} z_m^{j_{m-1} - j_t} \mathbb{1}(u_i < z_m)$$
$$+ \frac{z_m^{j_{m-1}}}{\prod_{t=0}^{m-1}} \mathbb{1}(u_i > z_m),$$

where $j_0 = k$. When $|K| = 1$, we have

$$\begin{aligned}&\Pr\left(U_{k,i} \le u_i, e_k = 1, e_{j_1} = 1\right)\\&= \int_0^{\min(u_i,z_1)} z^{k-1}\mathrm{d}z = \frac{u_i^k}{k}\mathbb{1}\left(u_i \le z_1\right) + \frac{z_1^k}{k}\mathbb{1}\left(u_i > z_1\right).\end{aligned}$$

Now, let us suppose the claim it is true for $|K| = m-1$, i.e.

$$\begin{aligned}A_{m-1} = &\sum_{t=0}^{m-2} (-1)^t \frac{u_i^{j_t}}{\prod_{r=0}^{t-1}(j_t - j_r) j_t \prod_{r=t+1}^{m-2}(j_r - j_t)} z_{m-1}^{j_{m-2}-j_t}\mathbb{1}(u_i < z_{m-1})\\&+ \frac{z_{m-1}^{j_{m-2}}}{\prod_{t=0}^{m-2}}\mathbb{1}(u_i > z_{m-1}),\end{aligned}$$

and prove it for $|K| = m$.

$$\begin{aligned}A_m &= \int_0^{z_m} z_{m-1}^{j_{m-1}-j_{m-2}-1} A_{m-1}\mathrm{d}z_{m-1}\\&= \sum_{t=0}^{m-2} (-1)^t \frac{u_i^{j_t}}{\prod_{r=0}^{t-1}(j_t - j_r) j_t \prod_{r=t+1}^{m-2}(j_r - j_t)} \frac{z_m^{j_{m-1}-j_t} - u_i^{j_{m-1}-j_t}}{j_{m-1} - j_t}\mathbb{1}(u_i < z_m)\\&+ \frac{1}{\prod_{t=0}^{m-2}}\left(\int_0^{u_i} z_{m-1}\mathrm{d}z_{m-1}\mathbb{1}(u_i < z_m) + \int_0^{z_m} z_{m-1}\mathrm{d}z_{m-1}\mathbb{1}(u_i > z_m)\right)\end{aligned}$$

which proves the claim. By considering $z_m = 1$ and by noting that $\prod_{r=0}^{t-1}(j_t - j_r) = (-1)^t \prod_{r=0}^{t-1}(j_r - j_t)$ and substituting with $u_i = F_i(x_i)$, the proof is complete. □

3.3 Complete Record Times

In this section we derive some results on record times. Let

$$T(n) := \inf\left\{m \in \mathbb{N} : \sum_{i=1}^{m} R_i^{\mathrm{CR}} = n\right\}, \quad n \ge 2, \quad T(1) := 1, \tag{3.7}$$

be the arrival time of the n-th complete record, where $\inf \emptyset := \infty$, which describes the case when there is no further complete record. We have seen in equation (3.2) that the number of complete records is finite with probability one, if the dimension d of our observations is at least 2. We start with the exact distribution of $T(n)$. Note that the distribution of $T(n)$ does not depend on the underlying univariate df $F_1, \dots, F_d$, provided that they are continuous.

Proposition 3.3.1. *For a generic size $d \geq 2$ we have*

$$\Pr(T(n) = k) = k^{-d} \sum_{A \subseteq \{2,\ldots,k-1\}, |A| = n-2} \prod_{q \in A} q^{-d} \prod_{m \in A^c} \left(1 - m^{-d}\right), \qquad k \geq n, \quad (3.8)$$

where $|A|$ is the total number of elements in the set A.

Proof. Note that

$$\begin{aligned} \Pr(T(n) = k) &= \Pr\left(R_n^{\text{CR}} = 1, S_{k-1} = n-1\right) \\ &= \Pr\left(R_n^{\text{CR}} = 1\right) \Pr\left(S_{k-1} = n-1\right), \end{aligned}$$

where $S_{k-1} = \sum_{m=1}^{k-1} R_m^{\text{CR}}$ is a sum of independent Bernoulli random variables, each with parameter m^{-d}. Therefore,

$$\Pr\left(S_{k-1} = n-1\right) = \Pr\left(\sum_{m=2}^{k-1} R_m^{\text{CR}} = n-2\right) = \sum_{A \in \mathcal{A}_{n-2}} \prod_{q \in A} q^{-d} \prod_{m \in A^c} \left(1 - m^{-d}\right),$$

which is a Poisson-Binomial distribution. □

Example 3.3.2. For $n = 2$ we get

$$\Pr(T(2) = k) = \frac{1}{k^d} \prod_{j=2}^{k-1} \left(1 - \frac{1}{j^d}\right), \qquad k \geq 2$$

while if $n = 3$

$$\Pr(T(3) = k) = \frac{1}{k^d} \prod_{j=2}^{k-1} \left(1 - \frac{1}{j^d}\right) \sum_{i=2}^{k-1} \frac{1}{i^d - 1}, \qquad k \geq 3.$$

Thus in the special case $d = 2$ we obtain

$$\Pr(T(2) = k) = \frac{1}{2k(k-1)}, \qquad \Pr(T(3) = k) = \frac{3k^2 - 7k + 2}{8k^2(k-1)^2}. \qquad (3.9)$$

The sequence $T(n)$, $n \geq 2$, is a Markov chain, as it is in the univariate case, see, e.g., Galambos 1987, (Section 6.3). Note that the state space is now $\{2, 3, \ldots\} \cup \{\infty\}$.

Proposition 3.3.3. *The sequence $T(n)$, $n \geq 2$ forms a Markov chain with the following transition probabilities*

$$\Pr\left(T(n)=k|T(n-1)=j\right)=\begin{cases} k^{-d}, & for\ k=j+1,\\ k^{-d}\prod_{m=j+1}^{k-1}\left(1-m^{-d}\right), & for\ k>j+1,\\ \prod_{m=j+1}^{\infty}\left(1-m^{-d}\right), & for\ k=\infty>j, \end{cases} \tag{3.10}$$

with $j \geq n-1$. The state $\{\infty\}$ is absorbing, that is $\Pr(T(n)=\infty \mid T(n-1)=\infty)=1$, *when $n \geq 3$.*

Proof. For a finite sequence of finite states, by the independence of $R_1^{\mathrm{CR}}, R_2^{\mathrm{CR}}, \ldots$ we have

$$\begin{aligned} &\Pr\left(T(m)=j_m, 2\leq m\leq n\right)\\ &=\prod_{m=2}^{n}\Pr\left(R_{j_m}^{\mathrm{CR}}=1\right)\Pr\left(\sum_{i=j_{m-1}+1}^{j_m-1}R_i^{\mathrm{CR}}=0\right), \end{aligned} \tag{3.11}$$

where $j_1=2$ by convention. Using this formula we obtain for the conditional probability

$$\begin{aligned} &\Pr\left(T(n)=j_n|T(m)=j_m, 2\leq m\leq n-1\right)\\ &=\Pr\left(R_{j_n}^{\mathrm{CR}}=1\right)\Pr\left(\sum_{i=j_{n-1}+1}^{j_n-1}R_i^{\mathrm{CR}}=0\right). \end{aligned} \tag{3.12}$$

Note that

$$\Pr\left(T(n-1)=j_{n-1}\right)=\Pr\left(R_{j_{n-1}}^{\mathrm{CR}}=1\right)\Pr\left(\sum_{i=2}^{j_{n-1}-1}R_i^{\mathrm{CR}}=n-2\right)$$

and

$$\begin{aligned} \Pr\left(T(n)=j_n, T(n-1)=j_{n-1}\right)=&\Pr\left(R_{j_n}^{\mathrm{CR}}=1\right)\Pr\left(\sum_{i=j_{n-1}+1}^{j_n-1}R_i^{\mathrm{CR}}=0\right)\\ &\Pr\left(R_{j_{n-1}}^{\mathrm{CR}}=1\right)\Pr\left(\sum_{i=2}^{j_{n-1}-1}R_i^{\mathrm{CR}}=n-2\right). \end{aligned}$$

Therefore $\Pr(T(n) = j_n | T(n-1) = j_{n-1})$ is equal to right hand-side of (3.12).

For the case that a time moves from a finite state to infinity we have

$$\Pr(T(n) = \infty, T(m) = j_m, 2 \le m \le n-1) = \prod_{m=2}^{n-1} \Pr\left(R_{j_m}^{\text{CR}} = 1\right) \times \Pr\left(\sum_{i=j_{m-1}+1}^{j_m-1} R_i^{\text{CR}} = 0\right) \times \Pr\left(\sum_{i=j_m+1}^{\infty} R_i^{\text{CR}} = 0\right).$$

Using this result and the one in (3.11) we obtain

$$\Pr(T(n) = \infty | T(m) = j_m, 2 \le m \le n-1) = \Pr\left(\sum_{i=j_{n-1}+1}^{\infty} R_i^{\text{CR}} = 0\right). \tag{3.13}$$

Now, noting that $\Pr(T(n-1) = j_{n-1}) = \Pr\left(R_{j_{n-1}}^{\text{CR}} = 1\right)$ and

$$\Pr(T(n) = \infty, T(n-1) = j_{n-1}) = \Pr\left(R_{j_{n-1}}^{\text{CR}} = 1\right) \Pr\left(\sum_{i=j_{n-1}+1}^{\infty} R_i^{\text{CR}} = 0\right),$$

then $\Pr(T(n) = \infty | T(n-1) = j_{n-1})$ is equal to right hand-side of (3.13).

To compute the transition probabilities, note that $\Pr(R_n^{\text{CR}} = 1) = n^{-d}$ and $\Pr(R_n^{\text{CR}} = 0) = 1 - n^{-d}$. Finally, to complete the proof we need to check that

$$p_{n|n-1} = \sum_{k \ge j+1} \Pr(T(n) = k | T(n-1) = j) + \Pr(T(n) = \infty | T(n-1) = j) = 1,$$

for each $j \ge 2$, i.e.

$$(j+1)^{-d} + \sum_{k=j+2}^{\infty} k^{-d} \prod_{m=j+1}^{k-1} (1 - m^{-d}) + \prod_{m=j+1}^{\infty} (1 - m^{-d}) = 1.$$

We prove by induction that

$$\sum_{k=j+2}^{M} k^{-d} \prod_{m=j+1}^{k-1} (1 - m^{-d}) + \prod_{m=j+1}^{M} (1 - m^{-d}) = 1 - (j+1)^{-d}, \tag{3.14}$$

for each $M \in \mathbb{N}$.

Step $M = j + 2$:

$$(j+2)^{-d}(1-(j+1)^{-d}) + (1-(j+1)^{-d})(1-(j+2)^{-d}) = 1-(j+1)^{-d}.$$

Now, we suppose (3.14) is true for M and we prove it is true also for $M+1$.

$$\begin{aligned}
&\sum_{k=j+2}^{M+1} k^{-d} \prod_{m=j+1}^{k-1}(1-m^{-d}) + \prod_{m=j+1}^{M+1}(1-m^{-d}) \\
&= \sum_{k=j+2}^{M} k^{-d} \prod_{m=j+1}^{k-1}(1-m^{-d}) + (M+1)^{-d} \prod_{m=j+1}^{M}(1-m^{-d}) \\
&\quad + (M+1)^{-d} \prod_{m=j+1}^{M}(1-m^{-d}) \\
&= (1-(j+1)^{-d})((M+1)^{-d} + 1 - (M+1)^{-d}) = 1-(j+1)^{-d}.
\end{aligned}$$

Therefore $p_{n|n-1} = 1$ and the proof is completed.

□

Chapter 4

Records for Stationary Dependent Sequences

4.1 Introduction

Let $\{X_n, n \geq 1\}$ be a sequence of identically distributed random variables (rvs), and denote by F the common univariate marginal distribution function. Throughout this chapter, we use the following notation for maxima: For any $i, j \in \mathbb{N}$, set $M_{i:j} := \max(X_i, \ldots, X_j)$. For simplicity, we set $M_j := M_{1:j}$, that is $M_j := \max(X_1, \ldots, X_j)$.

Except for Haiman (1987), Haiman et al. (1998), Ballerini and Resnick (1987), Arnold et al. (1998) as far as we know, most of the available results on records concern sequences of independent random variables or vectors. In the present work we derive some new results on the records of a stationary sequence of dependent random variables and dependent random vectors, under appropriate conditions of the dependence structure.

At first we consider a univariate second-order stationary Gaussian process with zero-mean, unit-variance. This means that for every $n = 1, 2, \ldots$, $\mathrm{E}(X_n) = 0$, $\mathrm{E}^2(X_n) = 1$ and the autocovariance of the process is translation-invariant depending only on the time difference, i.e. for every i, j, $\rho_{i,j} = \mathrm{E}(X_i X_j) = \mathrm{E}(X_0 X_{j-i}) = \rho_{0,j-i} \equiv \rho_{j-i}$, where ρ_{j-i} is a function only of the separation $j - i$ and for every m,

$\rho_{i+m,j+m} = \rho_{j-i}$. We derive the probability that a record at time n, say X_n, takes place, and the distribution of X_n, being a record. Furthermore, we derive the joint distribution of the arrival time process of records and more specifically the distribution of the increments between the first and second record and the third and second record. We compute the expected number of records which, depending on the type of correlation structure of the Gaussian process, can be finite or infinite. We also focus on joint records and we derive the probability that two consecutive and non-consecutive records at the time j and n, say X_j and X_n, take place, as well as the joint distribution of (X_j, X_n), considering they are both records.

We highlight that many of our findings, such as the probability that the records X_n and (X_j, X_n) take place and the arrival time of the n-th record, are independent of the marginal distribution function F, provided that is is continuous. As a consequence, the results actually hold for second-order stationary sequences with *Gaussian copulas*. On the contrary, the distribution of a record (two records), conditional to the assumption that it is a record (they are records), however does depend on F.

Next we consider a strictly stationary process satisfying some mild conditions on the tail behavior of the common marginal distribution function F and the long-range dependence of the extremes of the process. More specifically, it is assumed that F is attracted by the so-called Generalized Extreme-Value family of distributions, and that maxima on separated enough intervals within the time span n are approximately independent. Within this setting we derive the probability that X_n is a record, the distribution of X_n (being a record), and the expected number of records.

We complete the work by considering a zero-mean, unit-variance multivariate second-order stationary Gaussian process. We derive the probability that a complete record at time n occurs, and we compute the distribution of $\boldsymbol{X}_n$ (being a record), as well as the probability that two complete records at the time j and n occur, and the joint distribution of $(\boldsymbol{X}_j, \boldsymbol{X}_n)$ (being records).

The chapter is organized as follows. In Section 4.2.1 we introduce some notation and we briefly review some basic concepts on the multivariate closed skew-normal distribution. In Section 4.2.2 we present our main results on records for an univariate second-order stationary Gaussian process. In Section 4.2.3 we provide the asymptotic probability and distribution function of a record at time n for a strictly stationary process that satisfies some appropriate conditions. Finally, in Section 4.3 we extend some of the results derived in Section 4.2.3 to the case of multivariate second-order stationary Gaussian processes.

4.2 Univariate Case

4.2.1 Preliminary results and notation

Throughout the chapter we use the following notation. The symbol $\boldsymbol{X} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $n \in \mathbb{N}$, means an n-dimensional random vector that follows a multivariate Gaussian distribution with mean $\boldsymbol{\mu} \in \mathbb{R}^n$ and positive-definite covariance matrix $\boldsymbol{\Sigma} = \boldsymbol{\sigma}\bar{\boldsymbol{\Sigma}}\boldsymbol{\sigma} \in \mathbb{R}^{n,n}$, $\boldsymbol{\sigma} := \operatorname{diag}(\sigma_{11}, \ldots, \sigma_{nn})$, and $\bar{\boldsymbol{\Sigma}}$ is the correlation matrix. Its cumulative distribution function (cdf) and probability density function (pdf) are denoted by $\Phi_n(\boldsymbol{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ and $\phi_n(\boldsymbol{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ with $\boldsymbol{x} \in \mathbb{R}^n$. When $\boldsymbol{\mu} = \boldsymbol{0} = (0, \ldots, 0)^\top$ and $\boldsymbol{\Sigma} = \boldsymbol{I}$, where $\boldsymbol{I}$ is the identity matrix, we write $\Phi_n(\boldsymbol{x})$ for simplicity.

We indicate with $\boldsymbol{1}_{a,b}$ ($\boldsymbol{0}_{a,b}$) a matrix of dimension $a \times b$ whose elements are all equal to one (zero). We omit the subscripts when the dimensions of the matrices are clear from the context.

We introduce the notion of a multivariate *closed skew-normal* (CSN) random vector and we do so by using the so-called conditioning representation (Genton 2004, Ch. 2). Let $\boldsymbol{U} \sim N_m(\boldsymbol{\xi}, \boldsymbol{\Omega})$ being independent of $\boldsymbol{V} \sim N_n(\boldsymbol{0}, \boldsymbol{\Sigma})$, where $\boldsymbol{\xi} \in \mathbb{R}^m$, $\boldsymbol{\Omega} \in \mathbb{R}^m \times \mathbb{R}^m$ and $\boldsymbol{\Sigma} \in \mathbb{R}^n \times \mathbb{R}^n$. Let $\boldsymbol{\Delta} \in \mathbb{R}^n \times \mathbb{R}^m$, then

$$\begin{pmatrix} \boldsymbol{U} \\ \boldsymbol{\Delta}\boldsymbol{U} + \boldsymbol{V} \end{pmatrix} \sim N_{m+n}\left(\begin{pmatrix} \boldsymbol{\xi} \\ \boldsymbol{0} \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Omega} & \boldsymbol{\Omega}\boldsymbol{\Delta}^\top \\ \boldsymbol{\Delta}\boldsymbol{\Omega} & \boldsymbol{\Gamma} \end{pmatrix} \right),$$

where $\boldsymbol{\Gamma} = \boldsymbol{\Sigma} + \boldsymbol{\Delta}\boldsymbol{\Omega}\boldsymbol{\Delta}^\top$. Define $\boldsymbol{X}$ equal to $\boldsymbol{U}$, under the condition that $\boldsymbol{\Delta}\boldsymbol{U} + \boldsymbol{V} > \boldsymbol{\mu}$, denoted by $\boldsymbol{X} = (\boldsymbol{U} | \boldsymbol{\Delta}\boldsymbol{U} + \boldsymbol{V} > \boldsymbol{\mu})$, where $\boldsymbol{\mu} \in \mathbb{R}^n$. The m-dimensional random vector $\boldsymbol{X}$ follows a multivariate closed skew-normal distribution, in symbols

$\boldsymbol{X} \sim \text{CSN}_{m,n}(\boldsymbol{\xi}, \boldsymbol{\Omega}, \boldsymbol{\Delta}, \boldsymbol{\mu}, \boldsymbol{\Sigma})$, whose pdf is, for all $\boldsymbol{x} \in \mathbb{R}^m$,

$$\psi_{m,n}(\boldsymbol{x}; \boldsymbol{\xi}, \boldsymbol{\Omega}, \boldsymbol{\Delta}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{\phi_m(\boldsymbol{x} - \boldsymbol{\xi}; \boldsymbol{\Omega})\Phi_n(\boldsymbol{\Delta}(\boldsymbol{x} - \boldsymbol{\xi}); \boldsymbol{\mu}, \boldsymbol{\Sigma})}{\Phi_n(\boldsymbol{0}; \boldsymbol{\mu}, \boldsymbol{\Gamma})}. \tag{4.1}$$

We denote the cdf of $\boldsymbol{X}$ by $\Psi_{m,n}(\boldsymbol{x}; \boldsymbol{\xi}, \boldsymbol{\Omega}, \boldsymbol{\Delta}, \boldsymbol{\mu}, \boldsymbol{\Sigma})$. When $\boldsymbol{\xi} = \boldsymbol{0}$, $\boldsymbol{\Omega} = \boldsymbol{I}$ and $\boldsymbol{\mu} = \boldsymbol{0}$, we omit them among the parameters for simplicity and we write $\Psi_{m,n}(\boldsymbol{x}; \boldsymbol{\Delta}, \boldsymbol{\Sigma})$ and $\psi_{m,n}(\boldsymbol{x}; \boldsymbol{\Delta}, \boldsymbol{\Sigma})$ instead. We recall that the closed skew-normal distribution is also known in the literature as the unified multivariate skew-normal distribution, which simply uses a different parametrization (e.g, Ch. 7.1.2 in Azzalini 2013). The exposition of our results benefits from the parametrization used by the closed skew-normal distribution.

We recall that if $\boldsymbol{X} \sim \text{CSN}_{m,n}(\boldsymbol{\xi}, \boldsymbol{\Omega}, \boldsymbol{\Delta}, \boldsymbol{\mu}, \boldsymbol{\Sigma})$ then

$$\Psi_{m,n}(\boldsymbol{x}; \boldsymbol{\xi}, \boldsymbol{\Omega}, \boldsymbol{\Delta}, \boldsymbol{\Sigma}) = \frac{\Phi_{n+m}(\tilde{\boldsymbol{x}}; \tilde{\boldsymbol{\Omega}})}{\Phi_n(\boldsymbol{0}; \boldsymbol{\mu}, \boldsymbol{\Gamma})}, \tag{4.2}$$

where

$$\tilde{\boldsymbol{x}} = \begin{pmatrix} -\boldsymbol{\mu} \\ \boldsymbol{x} - \boldsymbol{\xi} \end{pmatrix}, \quad \tilde{\boldsymbol{\Omega}} = \begin{pmatrix} \boldsymbol{\Gamma} & -\boldsymbol{\Omega}\boldsymbol{\Delta}^\top \\ \boldsymbol{\Delta}\boldsymbol{\Omega} & \boldsymbol{\Omega} \end{pmatrix},$$

see Azzalini and Bacchieri (2010). Furthermore, for $\boldsymbol{b} \in \mathbb{R}^m$ and $\boldsymbol{A} \in \mathbb{R}^{q,m}$ then,

$$\boldsymbol{b} + \boldsymbol{X} \sim \text{CSN}_{m,n}(\boldsymbol{\xi} + \boldsymbol{b}, \boldsymbol{\Omega}, \boldsymbol{\Delta}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) \tag{4.3}$$

$$\boldsymbol{A}\boldsymbol{X} \sim \text{CSN}_{q,n}(\boldsymbol{A}\boldsymbol{\xi}, \boldsymbol{\Omega}^*, \boldsymbol{\Delta}^*, \boldsymbol{\mu}, \boldsymbol{\Sigma}^*) \tag{4.4}$$

where $\boldsymbol{\Omega}^* = \boldsymbol{A}\boldsymbol{\Omega}\boldsymbol{A}^\top$, $\boldsymbol{\Delta}^* = \boldsymbol{\Delta}\boldsymbol{\Omega}\boldsymbol{A}^\top\boldsymbol{\Omega}^{*-1}$ and $\boldsymbol{\Sigma}^* = \boldsymbol{\Gamma} - \boldsymbol{\Delta}^*\boldsymbol{A}\boldsymbol{\Omega}\boldsymbol{\Delta}^\top$, (see Ch. 2 in Genton 2004 for details).

4.2.2 Records of dependent univariate Gaussian sequences

Let $\{X_n, n \geq 1\}$ be a second-order stationary Gaussian sequence of dependent rvs. Without loss of generality, assume for simplicity that $\text{E}(X_i) = 0$, $\text{E}(X_i^2) = 1$ for every $1 \leq i \leq n$. From now on, we will refer to such a process as a stationary standard Gaussian (SSG) sequence. For any $n \in \mathbb{N}$, let $\mathcal{I} \subset \{1, \ldots, n\}$ and $\mathcal{I}^\complement = \{1, \ldots, n\} \backslash \mathcal{I}$ identify the $|\mathcal{I}|$-dimensional and $|\mathcal{I}^\complement|$-dimensional subvector partition such that $\boldsymbol{X} = (X_1, \ldots, X_n)^\top = (\boldsymbol{X}_{\mathcal{I}}^\top, \boldsymbol{X}_{\mathcal{I}^\complement}^\top)^\top$, with corresponding partition of the parameter $\bar{\boldsymbol{\Sigma}}$. By $|A|$ we denote the number of elements of a set A.

Our results rely on the following well-known important result on the conditional distribution derived from joint Gaussian distribution. Precisely, let $\boldsymbol{X} = (\boldsymbol{X}_{\mathcal{I}}^\top, \boldsymbol{X}_{\mathcal{I}^\complement}^\top)^\top \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with corresponding partition of the parameters $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, then in Anderson (1984, Theorem 2.5.1) it is established that the conditional distribution of $\boldsymbol{X}_{\mathcal{I}^\complement}$ given that $\boldsymbol{X}_{\mathcal{I}} = \boldsymbol{x}_{\mathcal{I}}$, is for all $\boldsymbol{x}_{\mathcal{I}} \in \mathbb{R}^{|\mathcal{I}|}$,

$$\begin{aligned} \boldsymbol{X}_{\mathcal{I}^\complement}|\boldsymbol{X}_{\mathcal{I}} = \boldsymbol{x}_{\mathcal{I}} &\sim N_{|\mathcal{I}^\complement|}\left(\boldsymbol{\mu}_{\mathcal{I}^\complement}, \boldsymbol{\Sigma}_{\mathcal{I}^\complement,\mathcal{I}^\complement;\mathcal{I}}\right), \\ \boldsymbol{\mu}_{\mathcal{I}^\complement} &= \boldsymbol{\Sigma}_{\mathcal{I}^\complement,\mathcal{I}}\bar{\boldsymbol{\Sigma}}_{\mathcal{I},\mathcal{I}}^{-1}\boldsymbol{x}_{\mathcal{I}}, \\ \boldsymbol{\Sigma}_{\mathcal{I}^\complement,\mathcal{I}^\complement;\mathcal{I}} &= \bar{\boldsymbol{\Sigma}}_{\mathcal{I}^\complement,\mathcal{I}^\complement} - \boldsymbol{\Sigma}_{\mathcal{I}^\complement,\mathcal{I}}\bar{\boldsymbol{\Sigma}}_{\mathcal{I},\mathcal{I}}^{-1}\boldsymbol{\Sigma}_{\mathcal{I},\mathcal{I}^\complement}. \end{aligned} \tag{4.5}$$

Furthermore, we denote the related correlation matrix by

$$\bar{\boldsymbol{\Sigma}}_{\mathcal{I}^\complement,\mathcal{I}^\complement;\mathcal{I}} = \boldsymbol{\sigma}_{\mathcal{I}^\complement,\mathcal{I}^\complement;\mathcal{I}}^{-1}\boldsymbol{\Sigma}_{\mathcal{I}^\complement,\mathcal{I}^\complement;\mathcal{I}}\boldsymbol{\sigma}_{\mathcal{I}^\complement,\mathcal{I}^\complement;\mathcal{I}}^{-1},$$

where $\boldsymbol{\sigma}_{\mathcal{I}^\complement,\mathcal{I}^\complement;\mathcal{I}} = \mathrm{diag}(\boldsymbol{\Sigma}_{\mathcal{I}^\complement,\mathcal{I}^\complement;\mathcal{I}})$. For any $j \in \{a, \ldots, b\}$, when $\mathcal{I} = \{j\}$ we simplify the notation writing X_j and $\boldsymbol{X}_{a:b\setminus j} = (X_a, \ldots, X_{j-1}, X_{j+1}, \ldots, X_b)^\top$. When $j = a$ or $j = b$ we further simplify the notation by $\boldsymbol{X}_{2:b} = (X_2, \ldots X_b)^\top$ and $\boldsymbol{X}_{1:b-1} = (X_1, \ldots, X_{b-1})^\top$.

In our first result we compute the probability that X_n is a record together with its distribution.

Proposition 4.2.1. *Let $\{X_n, n \geq 1\}$ be a SSG sequence of rvs. For every $n \geq 2$, let $\mathcal{I} = \{n\}$, $\mathcal{I}^\complement = \{1, \ldots, n-1\}$. Then, the probability that X_n is a record and the distribution of X_n, given that it is a record, are equal to*

$$\begin{aligned} \Pr(R_n = 1) &= \Phi_{n-1}(\boldsymbol{0}; \boldsymbol{\Gamma}_{1:n-1;1:n-1}) \\ \Pr(X_n \leq x | R_n = 1) &= \Psi_{1,n-1}\left(x; \boldsymbol{\varrho}_{1:n-1}, \bar{\boldsymbol{\Sigma}}_{1:n-1,1:n-1;n}\right), \end{aligned}$$

where $\boldsymbol{\Gamma}_{1:n-1;1:n-1}$ is a $(n-1) \times (n-1)$ variance-covariance matrix whose entries of the associated correlation matrix $\bar{\boldsymbol{\Gamma}}_{1:n-1;1:n-1}$ are

$$\gamma_{i,j;n} = \frac{1 + \rho_{i,j} - \rho_{i,n} - \rho_{j,n}}{2\sqrt{(1-\rho_{i,n})(1-\rho_{j,n})}}, \; i \neq n, j \neq n \tag{4.6}$$

and $\bar{\boldsymbol{\Sigma}}_{1:n-1,1:n-1;n}$ *is a* $(n-1)\times(n-1)$ *correlation matrix with entries*

$$\rho_{i,j;n} = \frac{\rho_{i,j} - \rho_{i,n}\rho_{j,n}}{\sqrt{(1-\rho_{i,n}^2)(1-\rho_{j,n}^2)}}, \; i \neq n, j \neq n.$$

Proof. The probability that X_n is a record is

$$\begin{aligned}
\Pr(X_n > M_{n-1}) &= \int_{-\infty}^{+\infty} \Pr\left(X_i < z, \forall i \in \mathcal{I}^{\complement} | X_n = z\right) \phi(z)\mathrm{d}z \\
&= \int_{-\infty}^{+\infty} \Pr(\boldsymbol{Z}_{1:n-1} \leq z\boldsymbol{\varrho}_{1:n-1})\phi(z)\mathrm{d}z \\
&= \mathrm{E}_Z\{\Pr(\boldsymbol{Z}_{1:n-1} \leq Z\boldsymbol{\varrho}_{1:n-1}|Z)\} \\
&= \Pr(\boldsymbol{Z}_{1:n-1} - Z\boldsymbol{\varrho}_{1:n-1} \leq \boldsymbol{0}) \equiv \Phi_{n-1}(\boldsymbol{0}; \boldsymbol{\Gamma}_{1:n-1;1:n-1}),
\end{aligned}$$

where

$$\begin{aligned}
\boldsymbol{\Gamma}_{1:n-1;1:n-1} &= \bar{\boldsymbol{\Sigma}}_{1:n-1,1:n-1;n} + \boldsymbol{\varrho}_{1:n-1}\boldsymbol{\varrho}_{1:n-1}^{\top} && (4.7)\\
\boldsymbol{\varrho}_{1:n-1} &= \boldsymbol{\sigma}_{1:n-1,1:n-1;n}^{-1}(\boldsymbol{1}_{n-1} - \bar{\boldsymbol{\Sigma}}_{1:n-1,n}) && (4.8)\\
&= \left(\sqrt{\frac{1-\rho_{i,n}}{1+\rho_{i,n}}}, \forall i \in \mathcal{I}^{\complement}\right)^{\top}.
\end{aligned}$$

To obtain the second line we used the formula in (4.5), which leads to $\boldsymbol{Z}_{1:n-1} = \boldsymbol{\sigma}_{1:n-1,1:n-1;n}^{-1}(\boldsymbol{X}_{1:n-1} - \boldsymbol{\mu}_n) \sim N_{n-1}(\boldsymbol{0}; \bar{\boldsymbol{\Sigma}}_{1:n-1,1:n-1;n})$, where $\boldsymbol{\mu}_n = (\rho_{i,n}, \forall i \in \mathcal{I}^{\complement})^{\top} v$, and this can be seen as independent of $Z \sim N(0,1)$. From the third to fourth row we used Lemma 7.1 in Azzalini and Valle (1996). With similar steps, we obtain the distribution for the record X_n,

$$\begin{aligned}
\Pr(X_n \leq x | R_n = 1) &= \frac{\Pr(X_n \leq x, X_n > M_{n-1})}{\Pr(X_n > M_{n-1})}, \\
&= \frac{\int_{-\infty}^{x} \phi(z)\Phi_{n-1}(z\boldsymbol{\varrho}_{1:n-1}; \bar{\boldsymbol{\Sigma}}_{1:n-1,1:n-1;n})\mathrm{d}z}{\Phi_{n-1}(\boldsymbol{0}; \boldsymbol{\Gamma}_{1:n-1;1:n-1})} \\
&\equiv \Psi_{1,n-1}\left(x; \boldsymbol{\varrho}_{1:n-1}, \bar{\boldsymbol{\Sigma}}_{1:n-1,1:n-1;n}\right).
\end{aligned}$$

□

The correlations $\rho_{i,j}$, $1 \leq i < j \leq n$, in Proposition 4.3.1 satisfy $-1 \leq \rho_{i,j;n} \leq 1$ Kurowicka and Cooke (2006) but they must also be such as to satisfy $-1 \leq \gamma_{i,j;n} \leq 1$

or

$$(\rho_{i,n}+\rho_{j,n}-1)-2\sqrt{(1-\rho_{i,n})(1-\rho_{j,n})}\leq \rho_{i,j}\leq(\rho_{i,n}+\rho_{j,n}-1)\\ +2\sqrt{(1-\rho_{i,n})(1-\rho_{j,n})}.$$

REMARK 4.2.2. *Assume in Proposition 4.3.1 that $\rho_{i,j}=0$ for all $1\leq i\neq j\leq n$. Then,*

$$\begin{aligned}\Pr(R_n=1)&=\Phi_{n-1}(\mathbf{0};\boldsymbol{I}_{n-1}+\mathbf{1}_{n-1}\mathbf{1}_{n-1}^\top)\\&=\mathrm{E}\left(\Phi_{n-1}(\mathbf{1}_{n-1}Z;\boldsymbol{I}_{n-1})\right)\\&=\int_{-\infty}^{+\infty}\Phi_{n-1}(\mathbf{1}_{n-1}z;\boldsymbol{I}_{n-1})\phi(z)\,dz=\int_{-\infty}^{+\infty}\Phi^{n-1}(z)\phi(z)\,dz=n^{-1},\end{aligned}$$

where $Z\sim N(0,1)$. As expected, we obtain the results in Galambos (1987) and Lemma 1.1 in Falk et al. (2018). Furthermore,

$$\begin{aligned}\Pr(X_n\leq x|R_n=1)&=\Psi_{1,n-1}(x;\mathbf{1}_{n-1},\boldsymbol{I}_{n-1})\\&=n\int_{-\infty}^{x}\Phi_{n-1}(\mathbf{1}_{n-1}z;\boldsymbol{I}_{n-1})\phi(z)\,dz\\&=n\int_{-\infty}^{x}\Phi^{n-1}(z)\phi(z)\,dz=\Phi(x)^n.\end{aligned}$$

Let

$$T(k):=\inf\left\{m\in\mathbb{N}:\sum_{i=1}^{m}R_i=k\right\},\quad k\geq 2,\quad T(1):=1,$$

be the arrival time of the k-th record.

Lemma 4.2.3. *Let $\{T(k)\}_{k\geq2}$ be the arrival time process of records. Let $\mathcal{I}=\{j_2,\dots,j_k\}$ where $2\leq j_2<\dots<j_k\in\mathbb{N}$ and $j_1:=1$. Set $\mathcal{I}^{\complement}:=\{1,\dots,j_k\}\setminus\mathcal{I}$. Then,*

$$\begin{aligned}&\Pr(T(i)=j_i,i=2,\dots,k)\\&=\Phi_{j_k-k}(\mathbf{0};\boldsymbol{\Gamma}_{\mathcal{I}^{\complement},\mathcal{I}^{\complement}})\Psi_{k-1,j_k-k}(\mathbf{0};\boldsymbol{D}\bar{\boldsymbol{\Sigma}}_{\mathcal{I},\mathcal{I}}\boldsymbol{D}^\top,\boldsymbol{\Delta},\bar{\boldsymbol{\Sigma}}_{\mathcal{I}^{\complement},\mathcal{I}^{\complement};\mathcal{I}})\end{aligned}$$

where $\boldsymbol{D}=(\boldsymbol{I}_{k-1}\quad \mathbf{0}_{k-1})-(\mathbf{0}_{k-1}\quad \boldsymbol{I}_{k-1})$,

$$\boldsymbol{\Delta}\;=\;\boldsymbol{\varrho}_{\mathcal{I}^{\complement},\mathcal{I}^{\complement}}\bar{\boldsymbol{\Sigma}}_{\mathcal{I},\mathcal{I}}\boldsymbol{D}^\top\left(\boldsymbol{D}\bar{\boldsymbol{\Sigma}}_{\mathcal{I},\mathcal{I}}\boldsymbol{D}^\top\right)^{-1}, \tag{4.9}$$

$$\begin{aligned}\boldsymbol{\Gamma}_{\mathcal{I}^\complement,\mathcal{I}^\complement} &= \boldsymbol{\varrho}_{\mathcal{I}^\complement,\mathcal{I}^\complement}\bar{\boldsymbol{\Sigma}}_{\mathcal{I},\mathcal{I}}\boldsymbol{\varrho}^\top_{\mathcal{I}^\complement,\mathcal{I}^\complement} + \bar{\boldsymbol{\Sigma}}_{\mathcal{I}^\complement,\mathcal{I}^\complement;\mathcal{I}}, \quad (4.10)\\ \boldsymbol{\varrho}_{\mathcal{I}^\complement,\mathcal{I}^\complement} &= \boldsymbol{\sigma}^{-1}_{\mathcal{I}^\complement,\mathcal{I}^\complement;\mathcal{I}}(\boldsymbol{B} - \boldsymbol{\Sigma}_{\mathcal{I}^\complement,\mathcal{I}}\bar{\boldsymbol{\Sigma}}^{-1}_{\mathcal{I},\mathcal{I}}), \quad (4.11)\end{aligned}$$

and

$$\boldsymbol{B} := \begin{pmatrix} \mathbf{1}_{j_2-2} & \mathbf{0}_{j_2-2} & \dots & \mathbf{0}_{j_2-2} & \mathbf{0}_{j_2-2} \\ \mathbf{0}_{j_3-j_2-1} & \mathbf{1}_{j_3-j_2-1} & \dots & \mathbf{0}_{j_3-j_2-1} & \mathbf{0}_{j_3-j_2-1} \\ \vdots & \vdots & & \vdots & \vdots \\ \mathbf{0}_{j_k-j_{k-1}-1} & \mathbf{0}_{j_k-j_{k-1}-1} & \dots & \mathbf{1}_{j_k-j_{k-1}-1} & \mathbf{0}_{j_k-j_{k-1}-1} \end{pmatrix} \in \mathbb{R}^{j_k-k,k-1} \quad (4.12)$$

Proof. We have

$$\begin{aligned} &\Pr(T(i) = j_i, i = 2, \dots, k) \\ &= \Pr(M_{j_i+1:j_{i+1}-1} < X_i, i = 1, \dots, k-1, X_{j_{k-1}} < X_{j_k}) \\ &= \int_{-\infty}^{+\infty}\int_{-\infty}^{z_k}\cdots\int_{-\infty}^{z_2} \Pr(M_{j_i+1:j_{i+1}-1} < z_i, i = 1, \dots, k-1 | X_{j_i} = z_i, i = 1, \dots, k-1) \\ &\quad \cdot \phi_k(z_1, \dots, z_k) \mathrm{d}z_1 \dots \mathrm{d}z_k \\ &= \int_{-\infty}^{+\infty}\int_{-\infty}^{z_k}\cdots\int_{-\infty}^{z_2} \Pr(\boldsymbol{X}_{\mathcal{I}^\complement} < \boldsymbol{B}\boldsymbol{z} | \boldsymbol{X}_{\mathcal{I}} = \boldsymbol{z})\phi_k(\boldsymbol{z}; \bar{\boldsymbol{\Sigma}}_{\mathcal{I},\mathcal{I}})\mathrm{d}\boldsymbol{z} \end{aligned}$$

where $\boldsymbol{B}$ is given in (4.12). By standardizing the random vector $\boldsymbol{X}_{\mathcal{I}^\complement}$, we obtain

$$\begin{aligned} &\int_{-\infty}^{+\infty}\int_{-\infty}^{z_k}\cdots\int_{-\infty}^{z_2} \Phi_{j_k-k}(\boldsymbol{\varrho}_{\mathcal{I}^\complement,\mathcal{I}^\complement}\boldsymbol{z}; \bar{\boldsymbol{\Sigma}}_{\mathcal{I}^\complement,\mathcal{I}^\complement;\mathcal{I}})\phi_k(\boldsymbol{z}; \bar{\boldsymbol{\Sigma}}_{\mathcal{I},\mathcal{I}})\mathrm{d}\boldsymbol{z} \\ &= \Phi_{j_k-k}(\mathbf{0}; \boldsymbol{\Gamma}_{\mathcal{I}^\complement,\mathcal{I}^\complement}) \int_{-\infty}^{+\infty}\int_{-\infty}^{z_k}\cdots\int_{-\infty}^{z_2} \psi_{k,j_k-k}(\mathbf{0}; \bar{\boldsymbol{\Sigma}}_{\mathcal{I},\mathcal{I}}, \boldsymbol{\varrho}_{\mathcal{I}^\complement,\mathcal{I}^\complement}, \bar{\boldsymbol{\Sigma}}_{\boldsymbol{X}_{\mathcal{I}^\complement},\boldsymbol{X}_{\mathcal{I}^\complement};\boldsymbol{X}_{\mathcal{I}}})\mathrm{d}\boldsymbol{z} \\ &= \Phi_{j_k-k}(\mathbf{0}; \boldsymbol{\Gamma}_{\mathcal{I}^\complement,\mathcal{I}^\complement}) \Pr(Z_1 < Z_2 < \dots < Z_k) \\ &= \Phi_{j_k-k}(\mathbf{0}; \boldsymbol{\Gamma}_{\mathcal{I}^\complement,\mathcal{I}^\complement}) \Pr(Z_1 - Z_2 < 0, \dots, Z_{k-1} - Z_k < 0) \end{aligned}$$

where $\boldsymbol{\Gamma}_{\mathcal{I}^\complement,\mathcal{I}^\complement}$ and $\boldsymbol{\varrho}_{\mathcal{I}^\complement,\mathcal{I}^\complement}$ are given in (4.10) and (4.11).

By recalling formula (4.4), we obtain

$$\begin{aligned} \begin{pmatrix} Z_1 - Z_2 \\ \vdots \\ Z_{k-1} - Z_k \end{pmatrix} &= \begin{pmatrix} 1 & -1 & 0 & 0 & \dots & 0 \\ 0 & 1 & -1 & 0 & \dots & 0 \\ \dots & & & & & \\ 0 & \dots & 0 & 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} Z_1 \\ \vdots \\ Z_k \end{pmatrix} \\ &= \boldsymbol{D}\boldsymbol{Z} \sim \mathrm{CSN}_{k-1,j_k-k}(\boldsymbol{D}\bar{\boldsymbol{\Sigma}}_{\mathcal{I},\mathcal{I}}\boldsymbol{D}^\top, \boldsymbol{\Delta}, \bar{\boldsymbol{\Sigma}}_{\mathcal{I}^\complement,\mathcal{I}^\complement;\mathcal{I}}) \end{aligned}$$

where $\mathbf{\Delta}$ is given in (4.9) $\square$

In the next result we establish the distribution of the arrival time $T(2)$ of the second record as well as that of the increment $X_{T(2)} - X_1$.

Theorem 4.2.4. *Let $\{X_n, n \geq 1\}$ be a SSG sequence of rvs. Let $\rho_{i,j} = \mathrm{E}(X_i, X_j)$ with $1 \leq i \neq j \leq n$. Assume that for $n \to \infty$, $\rho_{i,j} \to 0$ as $|j - i| \to \infty$ and $\rho_{k,n} \to 1$ as $k \to \infty$. For $n = 2, 3, \ldots$, the distribution of the arrival time of the second record $T(2)$ is*

$$\Pr\left(T(2) = n\right) = \begin{cases} 1/2, & n = 2, \\ \Phi_{n-2}(\mathbf{0}; \mathbf{\Gamma}_{2:n-1,2:n-1}) - \Phi_{n-1}(\mathbf{0}; \mathbf{\Gamma}_{2:n,2:n}), & n > 2 \end{cases} \tag{4.13}$$

where $\mathbf{\Gamma}_{2:n-1,2:n-1}$ and $\mathbf{\Gamma}_{2:n,2:n}$ are defined similarly to (4.7). Furthermore, for every $x > 0$, the distribution of the increment $X_{T(2)} - X_1$ is

$$H(x) = \sum_{n \geq 2} \Phi_{n-1}\left(\boldsymbol{u}_x; \mathbf{\Gamma}_{2:n,2:n}\right) - \Phi_{n-1}\left(\mathbf{0}; \mathbf{\Gamma}_{2:n,2:n}\right), \tag{4.14}$$

where $\boldsymbol{u}_x = (x/(1-\rho^2_{1,n})^{1/2}, 0, \ldots, 0)^\top$ is an $(n-1)$-dimensional vector.

Proof. When $n = 2$ we have

$$\Pr(T(2) = 2) = \Pr(X_2 > X_1) = 1/2.$$

For $n > 2$ we have

$$\begin{aligned} \Pr(T(2) = n) &= \Pr\left(X_i < X_1,\, i = 2, \ldots, n-1, X_n > X_1\right) \\ &= \Pr\left(X_i < X_1,\, i = 2, \ldots, n-1\right) - \Pr\left(X_i < X_1,\, i = 2, \ldots, n\right). \end{aligned}$$

Therefore, (4.13) follows by similar arguments to those used in Proposition 4.3.1. It must be checked that

$$\begin{aligned} \sum_{n \geq 2} \Pr(T(2) = n) &= \frac{1}{2} + \lim_{N \to \infty} \sum_{n=3}^{N} \left(\Phi_{n-2}(\mathbf{0}; \mathbf{\Gamma}_{2:n-1,2:n-1}) - \Phi_{n-1}(\mathbf{0}; \mathbf{\Gamma}_{2:n,2:n})\right) \\ &= \lim_{N \to \infty} \left(1 - \Phi_2\left(\mathbf{0}; \mathbf{\Gamma}_{1:2,1:2}\right) + \Phi_2\left(\mathbf{0}; \mathbf{\Gamma}_{1:2,1:2}\right)\right. \\ &\quad \left. - \cdots + \Phi_{N-2}\left(\mathbf{0}; \mathbf{\Gamma}_{1:N-2,1:N-2}\right) - \Phi_{N-1}\left(\mathbf{0}; \mathbf{\Gamma}_{1:N-1,1:N-1}\right)\right) \\ &= 1 - \lim_{N \to \infty} \Phi_{N-1}\left(\mathbf{0}; \mathbf{\Gamma}_{1:N-1,1:N-1}\right) = 1. \end{aligned}$$

Let $(\tilde{X}_1,\ldots,\tilde{X}_{n-1})$ be zero-mean unit-variance Gaussian sequence with variance-covariance matrix $\mathbf{\Gamma}_{1:n-1,1:n-1}$. Set

$$P_n = \Pr(\tilde{X}_i \leq 0,\ldots,\tilde{X}_{n-1} \leq 0) = \Phi_{n-1}\left(\mathbf{0};\mathbf{\Gamma}_{1:n-1,1:n-1}\right).$$

Clearly $\Phi_{n-1}\left(\mathbf{0};\mathbf{\Gamma}_{1:n-1,1:n-1}\right) = \Phi_{n-1}\left(\mathbf{0};\bar{\mathbf{\Gamma}}_{1:n-1,1:n-1}\right)$. We recall that $\Pr(\tilde{X}_i \leq 0) = 1/2$ for every $i = 1,\ldots,n-1$. By the Fréchet inequalities we have that

$$A_n := \max\left(0, \sum_{i=1}^{n} \Pr(X_i \leq 0) - (n-1)\right) = \max(0, 1-n/2) \leq P_n \leq 1/2.$$

For P_n we derive the following upper bound B_n. Precisely,

$$\begin{aligned}
P_n &= \Pr\left(\sum_{i=1}^{n-1} \mathbb{1}(\tilde{X}_i \leq 0) \geq n-1\right) = \Pr\left\{\sum_{i=1}^{n-1}\left(\mathbb{1}(\tilde{X}_i \leq 0) - \frac{1}{2}\right) \geq \frac{n-1}{2}\right\} \\
&\leq \Pr\left\{\left|\sum_{i=1}^{n-1}\left(\mathbb{1}(\tilde{X}_i \leq 0) - \frac{1}{2}\right)\right| \geq \frac{n-1}{2}\right\} \\
&\leq \frac{4}{(n-2)^2}\,\mathrm{E}\left[\left\{\sum_{i=1}^{n-1}\left(\mathbb{1}(\tilde{X}_i \leq 0) - \frac{1}{2}\right)\right\}^2\right] \\
&= \frac{4}{(n-2)^2}\sum_{i=1}^{n-1}\sum_{j=1}^{n-1} \mathrm{Cov}(\mathbb{1}(\tilde{X}_i \leq 0), \mathbb{1}(\tilde{X}_j \leq 0)) \\
&= \frac{4}{(n-2)^2}\sum_{i=1}^{n-1}\sum_{j=1}^{n-1} \mathrm{Cov}(P_{i,j;n} - 1/4) =: B_n,
\end{aligned}$$

where $P_{i,j;n} := \Pr(\tilde{X}_i \leq 0, \tilde{X}_j \leq 0) = \Phi_2(0;\gamma_{i,j;n})$ and where $\Phi_2(\cdot;\gamma_{i,j;n})$ is a bivariate Gaussian cdf with correlation $\gamma_{i,j;n}$ that is given in (4.6). In the third row we used the Chebyshev's inequality. Set $h = |j-i|$ we rewrite B_n as

$$\begin{aligned}
B_n &= \frac{4}{(n-2)^2}\sum_{h=0}^{n-2} 2(n-h)(P_{h;n} - 1/4) \\
&= \frac{8}{n(1+2/n)^2}(P_{0;n} - 1/4) + \frac{8}{n(1+2/n)^2}\sum_{h=1}^{n-2}\left(1 - \frac{h}{n}\right)(P_{h;n} - 1/4) \\
&= \alpha_n + \beta_n,
\end{aligned}$$

where $P_{h;n} := \Pr(\tilde{X}_0 \leq 0, \tilde{X}_h \leq 0) = \Phi_2(0; \gamma_{h;n})$ and

$$\gamma_{h;n} = \frac{1 + \rho_{0,h} - \rho_{0,n-i} - \rho_{h,n-i}}{2\sqrt{(1-\rho_{0,n-i})(1-\rho_{h,n-i})}}, \quad h = 0, \dots, n-2.$$

Now, when $h = 0$ we obtain $\gamma_{0;n} = 1$ and therefore $P_{0;n} = 1/2$ and as a consequence the term $\alpha_n \to 0$ as $n \to \infty$. We rewrite the term β_n as

$$\begin{aligned} \beta_n &= \frac{8}{n(1+2/n)^2} \sum_{h=1}^{n-2} (P_{h;n} - 1/4) - \frac{8}{n(1+2/n)^2} \sum_{h=1}^{n-2} \frac{h}{n} (P_{h;n} - 1/4) \\ &= c_n - d_n. \end{aligned}$$

Now by the assumption we have that for $n \to \infty$, $\gamma_{h;n} \to 0$ as $h \to \infty$, therefore for all $\varepsilon > 0$ there exists a n_0 such that for all $h > n_0$ we have $|P_{h;n} - 1/4| < \varepsilon$. As a consequence we have

$$\begin{aligned} c_n &= \frac{8}{n(1+2/n)^2} \left(\sum_{h=1}^{n_0} (P_{h;n} - 1/4) + \sum_{h=n_0+1}^{n-2} (P_{h;n} - 1/4) \right) \\ &< \frac{8}{n(1+2/n)^2} \left(c + \varepsilon(n - 2 + n_0 + 1) \right) = o(1), \end{aligned}$$

where c is a positive constant. Therefore, $c_n \to 0$ as $n \to \infty$ and since $d_n < c_n$ then $\beta_n \to 0$ and $B_n \to 0$ as $n \to \infty$. Concluding, since $A_n \leq P_n \leq B_n$ and $A_n = 0$ for $n \geq 2$, then $P_n \to 0$ as $n \to \infty$.

Finally, for every $x > 0$ the distribution of the increment $X_{T(2)} - X_1$ is

$$\begin{aligned} \Pr(X_{T(2)} - X_1 \leq x) &= \sum_{n \geq 2} \Pr(X_n - X_1 \leq x, T(2) = n) \\ &= \sum_{n \geq 2} \Pr(X_n - X_1 \leq x, X_i < X_1,\ i = 2, \dots, n-1, X_n > X_1) \\ &= \sum_{n \geq 2} \Pr(0 < X_n - X_1 \leq x, X_i < X_1,\ i = 2, \dots, n-1) \end{aligned}$$

The term inside the sum is equal to

$$\begin{aligned}
&\Pr(0 < X_n - X_1 \le x, X_i < X_1,\, i = 2, \dots, n-1) \\
&= \int_{-\infty}^{+\infty} \Pr(0 < X_n - u \le x, X_i < u,\, i = 2, \dots, n-1 | X_1 = z)\phi(z)\mathrm{d}z \\
&= \int_{-\infty}^{+\infty} \Pr(X_i < x,\, i = 2, \dots, n-1, X_n \le z + x | X_1 = z)\phi(z)\mathrm{d}z \\
&\quad - \int_{-\infty}^{+\infty} \Pr(X_i < z,\, i = 2, \dots, n | X_1 = z)\phi(z)\mathrm{d}z.
\end{aligned}$$

Therefore, (4.14) follows by similar arguments to those used in Proposition 4.3.1.

□

REMARK 4.2.5. *Note that when $\rho_{i,j} = 0$ for all $1 \le i < j \le n$ and $n > 2$ we obtain*

$$\begin{aligned}
\Pr(T(2) = n) &= \Phi_{n-2}(\mathbf{0}; \boldsymbol{I}_{n-2} + \mathbf{1}_{n-2}\mathbf{1}_{n-2}^\top) - \Phi_{n-1}(\mathbf{0}; \boldsymbol{I}_{n-1} + \mathbf{1}_{n-1}\mathbf{1}_{n-1}^\top) \\
&= \frac{1}{n-1} - \frac{1}{n} = \frac{1}{n(n-1)}.
\end{aligned}$$

Let $N := \sum_{n=1}^{\infty} R_n$ be the number of records among an infinite sequence $X_1, X_2, \dots$ When the components of the sequence are independent and identically distributed with a continuous df, then it is a well-known result that an infinite number of records will occur: $E(N) = \sum_{n=1}^{\infty} P(R_n = 1) = \sum_{n=1}^{\infty} 1/n = \infty$ Galambos (1987).

A natural question that arises is the following. What is the expected number of records that will take place in the case of a stationary Gaussian process?

Proposition 4.2.6. *Let $\{X_n\}_{n\ge1}$ be a SSG sequence of rvs and $\Phi_{n-1}(\mathbf{0}; \boldsymbol{\Gamma}_{1:n-1;1:n-1})$ be the probability that a record take place described in Proposition 4.3.1. Let N be the number of records among an infinite sequence $X_1, X_2, \dots$ Then, we have*

$$\mathrm{E}(N) = \begin{cases} \infty, & \text{if} \quad 1/2 \le \gamma_{i,j;n} \le 1, \quad \forall\, 1 \le i \ne j < n \\ 2, & \text{if} \quad \gamma_{i,j;n} = 0, \quad \forall\, 1 \le i \ne j < n. \end{cases}$$

where $\gamma_{i,j;n}$ is the correlation parameter in (4.6).

Proof. First, note that

$$\begin{aligned}
\mathrm{E}(N) = \mathrm{E}\left(\sum_{n=1}^{\infty} R_n\right) &= \sum_{n=1}^{\infty} \mathrm{E}(R_n) \\
&= 1 + \sum_{n=2}^{\infty} \Pr(X_n > M_{n-1}) \\
&= 1 + \sum_{n=2}^{\infty} \Phi_{n-1}(\mathbf{0}; \mathbf{\Gamma}_{1:n-1,1:n-1}).
\end{aligned}$$

The entries of the correlation matrix $\bar{\mathbf{\Gamma}}_{1:n-1,1:n-1}$ in (4.6) are $\gamma_{i,j;n} = 1/2$, $1 \le i \neq j < n$, if and only if $\rho_{i,j} = \rho_{i,n} = \rho_{j,n} = 0$. In this case by Remark 4.2.2 we have that $\Phi_{n-1}(\mathbf{0}; \boldsymbol{I}_{n-1} + \mathbf{1}_{n-1}\mathbf{1}_{n-1}^{\top}) = 1/n$. From this it follows that when $1/2 \le \gamma_{i,j;n} \le 1$ or

$$\sqrt{(1-\rho_{i,n})(1-\rho_{j,n})} \le 1 + \rho_{i,j} - \rho_{i,n} - \rho_{j,n} \le 2\sqrt{(1-\rho_{i,n})(1-\rho_{j,n})}, \qquad (4.15)$$

then $\Phi_{n-1}(\mathbf{0}; \mathbf{\Gamma}_{1:n-1,1:n-1}) \ge 1/n$ and as a consequence

$$\mathrm{E}(N) \ge \sum_{n=1}^{\infty} \frac{1}{n} = \infty.$$

For every $1 \le i \neq j < n$, provided that $\rho_{i,n} + \rho_{j,n} \ge 0$, when $\rho_{i,j} = \rho_{i,n} + \rho_{j,n} - 1$ then we have $\bar{\mathbf{\Gamma}}_{1:n-1,1:n-1} = \boldsymbol{I}_{n-1}$. Therefore in this case $\Phi_{n-1}(\mathbf{0}; \mathbf{\Gamma}_{1:n-1,1:n-1}) = \Phi_{n-1}(\boldsymbol{I}_{n-1}) = 2^{-n+1}$. As a consequence

$$\mathrm{E}(N) = 1 + \sum_{n=2}^{\infty} 2^{-n+1} = 2\sum_{n=0}^{\infty} 2^{-n} - 2 = 2.$$

□

From Proposition (4.2.6) it follows that the expected number of records depends on the type of correlation structure of the Gaussian process. For example, an infinite number of records is expected when all variables are uncorrelated or when X_i and X_j are more correlated than the sum of the correlations between X_i and X_n, and X_j and X_n, for every $1 \le i \neq j < n$. The second assertion follows from the left-hand side of the inequality in (4.15) by noting that $0 \le \sqrt{(1-\rho_{i,n})(1-\rho_{j,n})} \le 1$. This suggests looking at $1 + \rho_{i,j} - \rho_{i,n} - \rho_{j,n} \ge 1$ which holds as soon as

$\rho_{i,j} \geq \rho_{i,n} + \rho_{j,n}$. Instead, loosely speaking when X_i and X_j are less correlated than the sum of the correlations between X_i and X_n, and X_j and X_n, for every $1 \leq i \neq j < n$, the expected number of records can be finite. This assertion follows from the condition $\rho_{i,j} = \rho_{i,n} + \rho_{j,n} - 1$, provided that $\rho_{i,n} + \rho_{j,n} \geq 0$, which leads that two records should be expected.

In our next result we compute the distribution of the interarrival time between the second and third record.

Proposition 4.2.7. *The distribution of the increment has the representation*

$$\begin{aligned}&\Pr(X_{T(3)} - X_{T(2)} \leq x)\\ &= \sum_{j=2}^{\infty}\sum_{k=j+1}^{\infty} \Phi_{k-3}(\mathbf{0}; \mathbf{\Gamma}_{\mathcal{I}^{\complement},\mathcal{I}^{\complement}}) \left\{\Psi_{2,k-3}(\mathbf{0}; \boldsymbol{D}\bar{\mathbf{\Sigma}}_{\mathcal{I},\mathcal{I}}\boldsymbol{D}^{\top}, \mathbf{\Delta}, \bar{\mathbf{\Sigma}}_{\mathcal{I}^{\complement},\mathcal{I}^{\complement};\mathcal{I}})\right.\\ &\qquad\qquad \left. -\Psi_{2,k-3}((0,-x); \boldsymbol{D}\bar{\mathbf{\Sigma}}_{\mathcal{I},\mathcal{I}}\boldsymbol{D}^{\top}, \mathbf{\Delta}, \bar{\mathbf{\Sigma}}_{\mathcal{I}^{\complement},\mathcal{I}^{\complement};\mathcal{I}})\right\},\end{aligned} \tag{4.16}$$

where the sets of indices $\mathcal{I} = \{1, j, k\}$ *and* $\mathcal{I}^{\complement} = \{2, \ldots, j-1, j+1, \ldots, k-1\}$ *vary with* j *and* k, $\mathbf{\Delta}$ *and* $\tilde{\boldsymbol{\varrho}}_{\mathcal{I}^{\complement},\mathcal{I}^{\complement}}$ *are similarly defined as in formula* (4.9) *and* (4.11) *and where*

$$\boldsymbol{D} := \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix}$$

Proof. By the total probability rule

$$\Pr(X_{T(3)} - X_{T(2)} \leq x) = \sum_{j=2}^{\infty}\sum_{k=j+1}^{\infty} \Pr(X_k - X_j \leq x, T(3) = k, T(2) = j)$$

Note that, by repeating the same arguments as the previous proofs

$$\begin{aligned}&\Pr(X_k - X_j \leq x, T(3) = k, T(2) = j)\\ &= \Pr(X_k - X_j \leq x, M_{2:j-1} < X_1, X_1 < X_j, M_{j+1:k-1} < X_j, X_j < X_k)\\ &= \int_{-\infty}^{+\infty}\int_{z_k - x}^{z_k}\int_{-\infty}^{z_j} \Pr(M_{2:j-1} < z_1, M_{j+1:k-1} < z_j)\phi(z_1, z_j, z_k)\mathrm{d}z_1\mathrm{d}z_j\mathrm{d}z_k\\ &= \int_{-\infty}^{+\infty}\int_{z_k - x}^{z_k}\int_{-\infty}^{z_j} \Phi_{k-3}(\tilde{\boldsymbol{\varrho}}_{\mathcal{I}^{\complement},\mathcal{I}^{\complement}}\boldsymbol{z}; \bar{\mathbf{\Sigma}}_{\mathcal{I}^{\complement},\mathcal{I}^{\complement};\mathcal{I}})\phi_3(\boldsymbol{z}; \bar{\mathbf{\Sigma}}_{\mathcal{I},\mathcal{I}})\mathrm{d}\boldsymbol{z}\\ &= \Phi_{k-3}(\mathbf{0}; \mathbf{\Gamma}_{\mathcal{I}^{\complement},\mathcal{I}^{\complement}})\Pr(Z_1 < Z_j < Z_k) - \Phi_{k-3}(\mathbf{0}; \mathbf{\Gamma}_{\mathcal{I}^{\complement},\mathcal{I}^{\complement}})\Pr(Z_1 < Z_j < Z_k - x),\end{aligned}$$

and thus, the assertion follows by repeating the arguments in the proof of Lemma 4.2.3. □

In the following result we derive the probability that two records occur at prescribed indices, with no further record in between, together with the distribution of such consecutive records.

Theorem 4.2.8. *Let $\{X_n, n \geq 1\}$ be a SSG sequence of rvs. For every $n \geq 2$ and $j < n$, let $\mathcal{I} = \{j, n\}$, $\mathcal{I}^{\complement} = \{1, \dots, j-1, j+1, \dots, n-1\}$. The probability that two consecutive records X_j and X_n occur, is*

$$\begin{aligned} &\Pr\left(R_j = 1, R_n = 1, \cap_{i=j+1}^{n-1} R_i = 0\right) \\ &= \Phi_{n-2}(\mathbf{0}; \boldsymbol{\Gamma}_{1:n-1\setminus j,1:n-1\setminus j}) - \Phi_{n-1}(\mathbf{0}; \boldsymbol{\Gamma}_{1:n\setminus j,1:n\setminus j}), \end{aligned} \tag{4.17}$$

where $\boldsymbol{\Gamma}_{1:n-1\setminus j,1:n-1\setminus j}$ and $\boldsymbol{\Gamma}_{1:n\setminus j,1:n\setminus j}$ are similarly defined as in (4.7). The joint distribution of (X_j, X_n), given that they are consecutive records, is

$$\Pr\left(X_j \leq x_1, X_n \leq x_2 | R_j = 1, R_n = 1, \cap_{i=j+1}^{n-1} R_i = 0\right) = \begin{cases} P(x_1, x_2), & x_1 \leq x_2 \\ P(x_1, x_1) & x_1 > x_2 \end{cases}$$

where

$$\begin{aligned} P(a,b) = {}& w_{n-1}(b\boldsymbol{\mu}; \tilde{\boldsymbol{\Gamma}}_{1:n\setminus j,1:n\setminus j})\Psi_{1,n-1}\left(a; \tilde{\boldsymbol{\varrho}}_{1:n\setminus j}, -b\boldsymbol{\mu}, \bar{\boldsymbol{\Sigma}}_{1:n\setminus j,1:n\setminus j;j}\right) \\ & - w_{n-1}(\mathbf{0}; \boldsymbol{\Gamma}_{1:n\setminus j,1:n\setminus j})\Psi_{1,n-1}\left(a; \boldsymbol{\varrho}_{1:n\setminus j}, \mathbf{0}, \bar{\boldsymbol{\Sigma}}_{1:n\setminus j,1:n\setminus j;j}\right) \end{aligned}$$

and where $\boldsymbol{\varrho}_{1:n\setminus j}$ is similarly defined as in (4.8), $\tilde{\boldsymbol{\Gamma}}_{1:n\setminus j,1:n\setminus j} = \bar{\boldsymbol{\Sigma}}_{1:n\setminus j,1:n\setminus j;j} + \tilde{\boldsymbol{\varrho}}_{1:n\setminus j}\tilde{\boldsymbol{\varrho}}_{1:n\setminus j}^{\top}$ with

$$\tilde{\boldsymbol{\varrho}}_{1:n\setminus j} = \left(\boldsymbol{\varrho}_{1:n\setminus j}^{\top}, -\frac{\rho_{n,j}}{\sqrt{1-\rho_{n,j}^2}}\right)^{\top},$$

$$\boldsymbol{\mu} = \left(0, \dots, 0, (1-\rho_{j,n}^2)^{-1/2}\right)^{\top} \in \mathbb{R}^{n-1}.$$

and for any $\boldsymbol{x} \in \mathbb{R}^{n-1}$ and positive-definite matrix $\boldsymbol{\Sigma} \in \mathbb{R}^{n-1,n-1}$,

$$w_{n-1}(\boldsymbol{x}; \boldsymbol{\Sigma}) = \frac{\Phi_{n-1}(\boldsymbol{x}; \boldsymbol{\Sigma})}{\Phi_{n-2}(\mathbf{0}; \boldsymbol{\Gamma}_{1:n-1\setminus j,1:n-1\setminus j}) - \Phi_{n-1}(\mathbf{0}; \boldsymbol{\Gamma}_{1:n\setminus j,1:n\setminus j})}.$$

Proof. First we compute the probability that two consecutive records occur. For every $1 \leq j < n$ we have

$$\begin{aligned}\Pr(R_j = 1, R_n = 1, \cap_{i=j+1}^{n-1} R_i = 0) &= \Pr(X_j > M_{j-1}, X_n > M_{n-1}) \\ &= \Pr(X_i < X_j, \forall\, i \in \mathcal{I}^{\complement}, X_n > X_j) \\ &= \Pr(X_i < X_j, \forall\, i \in \mathcal{I}^{\complement},) \\ &- \Pr(X_i < X_j, \forall\, i \in \{i, \dots, n\} \setminus \{j\}).\end{aligned}$$

Therefore, (4.17) follows by similar arguments to those used in Proposition 4.3.1.

The joint distribution of (X_j, X_n) is given by

$$\begin{aligned}&\Pr\left(X_j \leq x_1, X_n \leq x_2 | R_j = 1, R_n = 1, \cap_{i=j+1}^{n-1} R_i = 0\right) \\ &= \frac{\Pr\left(X_j \leq x_1, X_n \leq x_2, X_j > M_{j-1}, X_n > M_{n-1}, \cap_{i=j+1}^{n-1} R_i = 0\right)}{\Pr(X_j > M_{j-1}, X_n > M_{n-1})}.\end{aligned}$$

Note that

$$\begin{aligned}&\Pr(X_j \leq x_1, X_n \leq x_2, X_j > M_{j-1}, X_n > M_{n-1}, \cap_{i=j+1}^{n-1} R_i = 0) \\ &= \Pr(X_j \leq x_1, X_n \leq x_2, X_i < X_j, \forall\, i \in \mathcal{I}^{\complement}, X_j < X_n) \\ &= \Pr(X_j \leq x_1, X_n \leq x_2, X_i < X_j, \forall\, i \in \mathcal{I}^{\complement}) \\ &- \Pr(X_j \leq x_1, X_n \leq x_2, X_i < X_j, i = 1, \dots, n, i \neq j) \\ &= A(x_1, x_2) - B(x_1, x_2).\end{aligned}$$

When $x_1 \leq x_2$, we obtain from similar arguments as those used in the proof of Proposition 4.3.1

$$\begin{aligned}A(x_1, x_2) &:= \Pr(X_j \leq x_1, X_n \leq x_2, X_i < X_j, \forall\, i \in \mathcal{I}^{\complement}) \\ &= \int_{-\infty}^{x_1} \Pr(X_n \leq x_2, X_i < z, \forall\, i \in \mathcal{I}^{\complement} | X_j = z)\phi(z)\mathrm{d}z \\ &= \int_{-\infty}^{x_1} \Pr\left(Z_i < \sqrt{\frac{1-\rho_{i,j}}{1+\rho_{i,j}}} z, \forall\, i \in \mathcal{I}^{\complement}, Z_n < \frac{x_2 - \rho_{n,j} z}{\sqrt{1-\rho_{n,j}^2}}\right)\phi(z)\mathrm{d}z \\ &= \int_{-\infty}^{x_1} \Phi_{n-1}(\tilde{\boldsymbol{\varrho}}_{1:n\setminus j} z + x_2 \boldsymbol{\mu}; \bar{\boldsymbol{\Sigma}}_{1:n\setminus j, 1:n\setminus j; j})\phi(z)\mathrm{d}z \\ &= \Phi_{n-1}(x_2 \boldsymbol{\mu}; \tilde{\boldsymbol{\Gamma}}_{1:n\setminus j, 1:n\setminus j})\Psi_{1,n-1}\left(x_1; \tilde{\boldsymbol{\varrho}}_{1:n\setminus j}, -x_2 \boldsymbol{\mu}, \bar{\boldsymbol{\Sigma}}_{1:n\setminus j, 1:n\setminus j; j}\right).\end{aligned}$$

Similarly,

$$\begin{aligned} B(x_1,x_2) &:= \Pr(X_j \le x_1, X_i < X_j, i=1,\dots,n, i\neq j) \\ &= \int_{-\infty}^{x_1} \Pr(X_i < z, i=1,\dots,n, i\neq j | X_j = z)\phi(z)\mathrm{d}z \\ &= \Phi_{n-1}(\mathbf{0};\mathbf{\Gamma}_{1:n\setminus j,1:n\setminus j})\Psi_{1,n-1}\left(x_1;\boldsymbol{\varrho}_{1:n\setminus j},\mathbf{0},\bar{\mathbf{\Sigma}}_{1:n\setminus j,1:n\setminus j;j}\right). \end{aligned}$$

When $x_1 > x_2$, it is sufficient to compute $A(x_1,x_2)$ and $B(x_1,x_2)$ in (x_2,x_2). □

In the next result we drop the assumption that the two records in Theorem 4.2.8 are consecutive.

Theorem 4.2.9. *Let $\{X_n, n\ge 1\}$ be a SSG sequence of rvs. For every $n \ge 2$ and $j < n$, let $\mathcal{I} = \{j,n\}$ and $\mathcal{I}^{\complement} = \{1,\dots,j-1,j+1,\dots,n-1\}$. The probability that X_j and X_n are records, is*

$$\Pr(R_j = 1, R_n = 1) = \Phi_{n-1}(\mathbf{0};\tilde{\mathbf{\Omega}}).$$

The joint distribution of (X_j, X_n), given that they are records, is

$$\Pr(X_j \le x_1, X_n \le x_2 | R_j = 1, R_n = 1) = \begin{cases} P(x_1,x_2), & x_1 \le x_2 \\ P(x_1,x_1) & x_1 > x_2 \end{cases}$$

where

$$\begin{aligned} P(a,b) = \frac{\Phi_{n-2}(\mathbf{0};\mathbf{\Gamma}_{\mathcal{I}^{\complement},\mathcal{I}^{\complement}})}{\Phi_{n-1}(\mathbf{0};\tilde{\mathbf{\Omega}})}\Big(&\Psi_{2,n-2}(a,b;\bar{\mathbf{\Sigma}}_{\mathcal{I},\mathcal{I}},\boldsymbol{\varrho}_{\mathcal{I}^{\complement},\mathcal{I}^{\complement}},\bar{\mathbf{\Sigma}}_{\mathcal{I}^{\complement},\mathcal{I}^{\complement};\mathcal{I}}) \\ &- \Psi_{2,n-2}(a,a;\bar{\mathbf{\Sigma}}_{\mathcal{I},\mathcal{I}},\boldsymbol{\varrho}_{\mathcal{I}^{\complement},\mathcal{I}^{\complement}},\bar{\mathbf{\Sigma}}_{\mathcal{I}^{\complement},\mathcal{I}^{\complement};\mathcal{I}}) \\ &+ \Psi_{2,n-2}(0,a;\boldsymbol{D}\bar{\mathbf{\Sigma}}_{\mathcal{I},\mathcal{I}}\boldsymbol{D}^{\top},\mathbf{\Delta},\bar{\mathbf{\Sigma}}^{**}_{\mathcal{I}^{\complement},\mathcal{I}^{\complement};\mathcal{I}})\Big). \end{aligned}$$

Proof. Similar steps as those used in the proof of Proposition 4.3.1 show that the probability that X_j and X_j are records, is

$$
\begin{aligned}
&\Pr(R_j = 1, R_n = 1) = \Pr(X_j > M_{j-1}, X_n > M_{n-1}) \\
&= \int_{-\infty}^{+\infty}\int_{-\infty}^{z_2} \Pr\left(M_{j-1} < z_1, M_{j+1}^{n-1} < z_2 | X_j = z_1, X_n = z_2\right) \phi_2(z_1, z_2; \bar{\boldsymbol{\Sigma}}_{\mathcal{I},\mathcal{I}}) \mathrm{d}z_1 \mathrm{d}z_2 \\
&= \int_{-\infty}^{+\infty}\int_{-\infty}^{z_2} \Phi_{n-2}\left(\boldsymbol{\varrho}_{\mathcal{I}^\complement,\mathcal{I}^\complement}\boldsymbol{z}; \bar{\boldsymbol{\Sigma}}_{\mathcal{I}^\complement,\mathcal{I}^\complement;\mathcal{I}}\right) \phi_2(\boldsymbol{z}; \bar{\boldsymbol{\Sigma}}_{\mathcal{I},\mathcal{I}}) \mathrm{d}z_1 \mathrm{d}z_2 \\
&= \Phi_{n-2}(\mathbf{0}, \boldsymbol{\Gamma}_{\mathcal{I},\mathcal{I}^\complement}) \int_{-\infty}^{+\infty}\int_{-\infty}^{z_2} \psi_{2,n-2}\left(\boldsymbol{z}; \bar{\boldsymbol{\Sigma}}_{\mathcal{I},\mathcal{I}}, \boldsymbol{\varrho}_{\mathcal{I}^\complement,\mathcal{I}^\complement}, \bar{\boldsymbol{\Sigma}}_{\mathcal{I}^\complement,\mathcal{I}^\complement;\mathcal{I}}\right) \mathrm{d}z_1 \mathrm{d}z_2 \\
&= \Phi_{n-2}(\mathbf{0}; \boldsymbol{\Gamma}_{\mathcal{I},\mathcal{I}^\complement}) \Pr(Z_1 - Z_2 < 0),
\end{aligned}
$$

where $(Z_1, Z_2) \sim \mathrm{CSN}_{2,n-2}(\bar{\boldsymbol{\Sigma}}_{\mathcal{I},\mathcal{I}}, \boldsymbol{\varrho}_{\mathcal{I}^\complement,\mathcal{I}^\complement}, \bar{\boldsymbol{\Sigma}}_{\mathcal{I}^\complement,\mathcal{I}^\complement;\mathcal{I}})$. Precisely, to obtain the third line we used the formula in (4.5) and where

$$
\boldsymbol{B} = \begin{pmatrix} \mathbf{1}_{j-1} & \mathbf{0}_{j-1} \\ \mathbf{0}_{n-j-1} & \mathbf{1}_{n-j-1} \end{pmatrix},
$$

$$
\boldsymbol{\varrho}_{\mathcal{I}^\complement,\mathcal{I}^\complement} = \boldsymbol{\sigma}^{-1}_{\mathcal{I}^\complement,\mathcal{I}^\complement;\mathcal{I}}(\boldsymbol{B} - \boldsymbol{\Sigma}_{\mathcal{I}^\complement,\mathcal{I}}\bar{\boldsymbol{\Sigma}}^{-1}_{\mathcal{I},\mathcal{I}})
$$

$$
= \begin{pmatrix}
\frac{1}{\sigma_{11}}\left(1 - \frac{\rho_{1j} - \rho_{1n}\rho_{jn}}{1-\rho_{jn}^2}\right) & \frac{\rho_{1j}\rho_{jn} - \rho_{1n}}{\sigma_{11}(1-\rho_{jn}^2)} \\
\vdots & \vdots \\
\frac{1}{\sigma_{j-1,j-1}}\left(1 - \frac{\rho_{j-1,j} - \rho_{j-1,n}\rho_{jn}}{1-\rho_{jn}^2}\right) & \frac{\rho_{j-1,j}\rho_{jn} - \rho_{j-1,n}}{\sigma_{j-1,j-1}(1-\rho_{jn}^2)} \\
\frac{\rho_{j+1,j}\rho_{jn} - \rho_{j+1,n}}{\sigma_{j+1,j+1}(1-\rho_{jn}^2)} & \frac{1}{\sigma_{j+1,j+1}}\left(1 - \frac{\rho_{j+1,j} - \rho_{j+1,n}\rho_{jn}}{1-\rho_{jn}^2}\right) \\
\vdots & \vdots \\
\frac{\rho_{n-1,j}\rho_{jn} - \rho_{n-1,n}}{\sigma_{n-1,n-1}(1-\rho_{jn}^2)} & \frac{1}{\sigma_{n-1,n-1}}\left(1 - \frac{\rho_{n-1,j} - \rho_{n-1,n}\rho_{jn}}{1-\rho_{jn}^2}\right)
\end{pmatrix} \tag{4.18}
$$

and $\bar{\boldsymbol{\Sigma}}_{\mathcal{I}^\complement,\mathcal{I}^\complement;\mathcal{I}}$ is a $(n-2)\times(n-2)$ partial correlation matrix with upper diagonal entries

$$
\rho_{i,k;j,n} = \rho_{ij} - \frac{\rho_{ij} - \rho_{in}\rho_{jn}}{1-\rho_{jn}^2}\rho_{kj} - \frac{\rho_{in} - \rho_{ij}\rho_{jn}}{1-\rho_{jn}^2}\rho_{kn}, \ \forall\, i < k \in \mathcal{I}^\complement.
$$

and

$$
\sigma_{i,i} = 1 - \frac{\rho_{ij} - \rho_{in}\rho_{jn}}{1-\rho_{jn}^2}\rho_{ij} - \frac{\rho_{in} - \rho_{ij}\rho_{jn}}{1-\rho_{jn}^2}\rho_{in}.
$$

In the third line we multiply and divide the term within the integrals with $\Phi_{n-2}(\mathbf{0}; \boldsymbol{\Gamma}_{\mathcal{I}^\complement,\mathcal{I}^\complement})$ where $\boldsymbol{\Gamma}_{\mathcal{I}^\complement,\mathcal{I}^\complement}$ is defined as

$$
\boldsymbol{\Gamma}_{\mathcal{I}^\complement,\mathcal{I}^\complement} = \boldsymbol{\varrho}_{\mathcal{I}^\complement,\mathcal{I}^\complement}\bar{\boldsymbol{\Sigma}}_{\mathcal{I},\mathcal{I}}\boldsymbol{\varrho}^\top_{\mathcal{I}^\complement,\mathcal{I}^\complement} + \bar{\boldsymbol{\Sigma}}_{\mathcal{I}^\complement,\mathcal{I}^\complement;\mathcal{I}}.
$$

We therefore recognize a unified multivariate skew-normal pdf within the integrals and the integral of it can be seen as $\Pr(Z_1 < Z_2)$. Now, by (4.4) we obtain

$$Z_1 - Z_2 = \begin{pmatrix} 1 & -1 \end{pmatrix} \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} \sim \mathrm{CSN}_{1,n-2}\left(2(1-\rho_{j,n}), \boldsymbol{\Delta}^*, \boldsymbol{\Sigma}^*_{\mathcal{I}^\complement,\mathcal{I}^\complement;\mathcal{I}}\right)$$

where

$$\begin{aligned}\boldsymbol{\Delta}^* &= \frac{1}{2(1-\rho_{j,n})} \boldsymbol{\varrho}_{\mathcal{I}^\complement,\mathcal{I}^\complement} \bar{\boldsymbol{\Sigma}}_{\mathcal{I},\mathcal{I}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} \mathbf{1}_{j-1} \\ -\mathbf{1}_{n-j-1} \end{pmatrix} \left(\frac{1}{\sigma_{ii}} \left(1 + \frac{\rho_{in} - \rho_{ij}}{1-\rho_{jn}} \right) \right)_{i=1,\dots,n-1, i\neq j}\end{aligned}$$

and

$$\boldsymbol{\Sigma}^*_{\mathcal{I}^\complement,\mathcal{I}^\complement;\mathcal{I}} = \boldsymbol{\Gamma}_{\mathcal{I}^\complement,\mathcal{I}^\complement} - \boldsymbol{\Delta}^* \begin{pmatrix} 1-\rho_{jn} & -1+\rho_{jn} \end{pmatrix} \boldsymbol{\varrho}^\top_{\mathcal{I}^\complement,\mathcal{I}^\complement}.$$

By formula (4.2) we obtain the result, with

$$\tilde{\boldsymbol{\Omega}} = \begin{pmatrix} \boldsymbol{\Gamma}_{\mathcal{I}^\complement,\mathcal{I}^\complement} & 2(1-\rho_{jn})\boldsymbol{\Delta}^{*\top} \\ 2(1-\rho_{jn})\boldsymbol{\Delta}^* & 2(1-\rho_{jn}) \end{pmatrix}.$$

By similar steps we can compute the joint distribution of two records (X_j, X_n) for $j < n$.

$$\begin{aligned}&\Pr(X_j \le x_1, X_n \le x_2 | R_j = 1, R_n = 1) \\ &= \frac{\Pr(X_j \le x_1, X_n \le x_2, X_j > M_{j-1}, X_n > M_{n-1})}{\Pr(X_j > M_{j-1}, X_n > M_{n-1})}.\end{aligned}$$

The numerator can be written as

$$\begin{aligned}
&\Pr(X_j \leq x_1, X_n \leq x_2, X_j > M_{j-1}, X_n > M_{n-1}) \\
&= \int_{-\infty}^{x_2}\int_{-\infty}^{\min(x_1,z_2)} \Pr\left(M_{j-1} < z_1, M_{j+1:n-1} < z_2 | X_j = z_1, X_n = z_2\right)\phi(\boldsymbol{z};\bar{\boldsymbol{\Sigma}}_{\mathcal{I},\mathcal{I}})\mathrm{d}z_1\mathrm{d}z_2 \\
&= \int_{-\infty}^{x_2}\int_{-\infty}^{x_1} \Pr\left(M_{j-1} < z_1, M_{j+1:n-1} < z_2 | X_j = z_1, X_n = z_2\right) \\
&\qquad \cdot \phi_2(\boldsymbol{z};\bar{\boldsymbol{\Sigma}}_{\mathcal{I},\mathcal{I}})\mathbb{1}(z_2 > x_1)\mathrm{d}z_1\mathrm{d}z_2 \\
&\quad + \int_{-\infty}^{x_2}\int_{-\infty}^{z_2} \Pr\left(M_{j-1} < z_1, M_{j+1:n-1} < z_2 | X_j = z_1, X_n = z_2\right) \\
&\qquad \cdot \phi_2(\boldsymbol{z};\bar{\boldsymbol{\Sigma}}_{\mathcal{I},\mathcal{I}})\mathbb{1}(z_2 < x_1)\mathrm{d}z_1\mathrm{d}z_2 \\
&= \int_{x_1}^{x_2}\int_{-\infty}^{x_1} \Pr\left(M_{j-1} < z_1, M_{j+1:n-1} < z_2 | X_j = z_1, X_n = z_2\right)\phi_2(\boldsymbol{z};\bar{\boldsymbol{\Sigma}}_{\mathcal{I},\mathcal{I}})\mathrm{d}z_1\mathrm{d}z_2 \\
&\quad + \int_{-\infty}^{x_1}\int_{-\infty}^{z_2} \Pr\left(M_{j-1} < z_1, M_{j+1:n-1} < z_2 | X_j = z_1, X_n = z_2\right)\phi_2(\boldsymbol{z};\bar{\boldsymbol{\Sigma}}_{\mathcal{I},\mathcal{I}})\mathrm{d}z_1\mathrm{d}z_2 \\
&= \int_{-\infty}^{x_2}\int_{-\infty}^{x_1} \Phi_{n-2}\left(\boldsymbol{\varrho}_{\mathcal{I}^\complement,\mathcal{I}^\complement}\boldsymbol{z};\bar{\boldsymbol{\Sigma}}_{\mathcal{I}^\complement,\mathcal{I}^\complement;\mathcal{I}}\right)\phi_2(\boldsymbol{z};\bar{\boldsymbol{\Sigma}}_{\mathcal{I},\mathcal{I}})\mathrm{d}z_1\mathrm{d}z_2 \\
&\quad - \int_{-\infty}^{x_1}\int_{-\infty}^{x_1} \Phi_{n-2}\left(\boldsymbol{\varrho}_{\mathcal{I}^\complement,\mathcal{I}^\complement}\boldsymbol{z};\bar{\boldsymbol{\Sigma}}_{\mathcal{I}^\complement,\mathcal{I}^\complement;\mathcal{I}}\right)\phi_2(\boldsymbol{z};\bar{\boldsymbol{\Sigma}}_{\mathcal{I},\mathcal{I}})\mathrm{d}z_1\mathrm{d}z_2 \\
&\quad + \int_{-\infty}^{x_1}\int_{-\infty}^{z_2} \Phi_{n-2}\left(\boldsymbol{\varrho}_{\mathcal{I}^\complement,\mathcal{I}^\complement}\boldsymbol{z};\bar{\boldsymbol{\Sigma}}_{\mathcal{I}^\complement,\mathcal{I}^\complement;\mathcal{I}}\right)\phi_2(\boldsymbol{z};\bar{\boldsymbol{\Sigma}}_{\mathcal{I},\mathcal{I}})\mathrm{d}z_1\mathrm{d}z_2
\end{aligned}$$

where $\boldsymbol{\varrho}_{\mathcal{I}^\complement,\mathcal{I}^\complement}$ is as in (4.18). We multiply and divide each term within the integrals with $\Phi_{n-2}(\mathbf{0};\boldsymbol{\Gamma}_{\mathcal{I}^\complement,\mathcal{I}^\complement})$. Then, we recognize that the first two integrals provide the distribution of the closed skew-normal random vector we introduced before, evaluated at the points (x_1,x_2), (x_1,x_1). Instead, the third integral represents the distribution of the random vector $(Z_1 - Z_2, Z_1)$ which again according to (4.4) follows a closed skew-normal distribution, i.e.,

$$\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} = \boldsymbol{D}\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} \sim \mathrm{CSN}_{2,n-2}(\boldsymbol{D}\bar{\boldsymbol{\Sigma}}_{\mathcal{I},\mathcal{I}}\boldsymbol{D}^\top, \boldsymbol{\Delta}, \bar{\boldsymbol{\Sigma}}^{**}_{\mathcal{I}^\complement,\mathcal{I}^\complement;\mathcal{I}}),$$

where $\boldsymbol{\Delta} = \boldsymbol{\varrho}_{\mathcal{I}^\complement,\mathcal{I}^\complement}\bar{\boldsymbol{\Sigma}}_{\mathcal{I}^\complement,\mathcal{I}^\complement}\boldsymbol{D}^\top(\boldsymbol{D}\bar{\boldsymbol{\Sigma}}_{\mathcal{I},\mathcal{I}}\boldsymbol{D}^\top)^{-1}$ and $\bar{\boldsymbol{\Sigma}}^{**}_{\mathcal{I}^\complement,\mathcal{I}^\complement;\mathcal{I}} = \boldsymbol{\Gamma}_{\mathcal{I}^\complement,\mathcal{I}^\complement} - \boldsymbol{\Delta}\boldsymbol{D}\bar{\boldsymbol{\Sigma}}_{\mathcal{I},\mathcal{I}}\boldsymbol{\varrho}^\top_{\mathcal{I}^\complement,\mathcal{I}^\complement}$.
□

It follows from Theorem 4.2.9 that the two events: a record occuring at time j and n, are not independent. Indeed, the probability $\Phi_{n-1}(\mathbf{0};\tilde{\boldsymbol{\Omega}})$ is different from the product of the two marginal probabilities $\Phi_{j-1}(\mathbf{0};\mathbf{\Gamma}_{1:j-1;1:j-1})$ and $\Phi_{n-1}(\mathbf{0};\mathbf{\Gamma}_{1:n-1;1:n-1})$, derived in Proposition 4.3.1.

REMARK 4.2.10. *The marginal distribution of X_j, given that (X_j, X_n) are records, is*

$$\begin{aligned}
&\Pr(X_j \leq x_1 | R_j = 1, R_n = 1)\\
&= \frac{\Phi_{n-2}(\mathbf{0};\mathbf{\Gamma}_{\mathcal{I}^\complement,\mathcal{I}^\complement})}{\Phi_{n-1}(\mathbf{0};\tilde{\mathbf{\Omega}})} \times \big(\Psi_{1,n-2}(x_1; 1, \mathbf{\Delta}_1, \bar{\mathbf{\Omega}}_{\mathcal{I}^\complement,\mathcal{I}^\complement;\mathcal{I}})\\
&\qquad\qquad - \Psi_{2,n-2}(x_1, x_1; \bar{\mathbf{\Sigma}}_{\mathcal{I},\mathcal{I}}, \boldsymbol{\varrho}_{\mathcal{I}^\complement,\mathcal{I}^\complement}, \bar{\mathbf{\Sigma}}_{\mathcal{I}^\complement,\mathcal{I}^\complement;\mathcal{I}})\\
&\qquad\qquad + \Psi_{2,n-2}(0, x_1; \boldsymbol{D}\bar{\mathbf{\Sigma}}_{\mathcal{I},\mathcal{I}}\boldsymbol{D}^\top, \mathbf{\Delta}, \bar{\mathbf{\Sigma}}_{\mathcal{I}^\complement,\mathcal{I}^\complement;\mathcal{I}})\big).
\end{aligned}$$

where

$$\mathbf{\Delta}_1 = \begin{pmatrix} \sigma_{ii}^{-1}(1-\rho_{ij})_{i=1,\ldots,j-1} \\ \sigma_{ii}^{-1}(\rho_{jn}-\rho_{in})_{i=j+1,\ldots,n-1} \end{pmatrix},$$

$$\bar{\mathbf{\Omega}}_{\mathcal{I}^\complement,\mathcal{I}^\complement;\mathcal{I}} = \bar{\mathbf{\Sigma}}_{\mathcal{I}^\complement,\mathcal{I}^\complement;\mathcal{I}} + \begin{pmatrix} \sigma_{ii}^{-1}(\rho_{ij}\rho_{jn}-\rho_{in})_{i=1,\ldots,j-1} \\ \sigma_{ii}^{-1}(1-\rho_{jn}^2-\rho_{ij}-\rho_{in}\rho_{jn})_{i=j+1,\ldots,n-1} \end{pmatrix},$$

and these parameters are obtained from (4.4)*, with $\boldsymbol{A} := (0 \quad 1)$. See the proof of Theorem 4.2.9 for the details. Hence, similarly to the case of independent random variables in Falk et al. (2018), the distribution of X_j being a record is affected, if we know that X_n is a record as well. The marginal distribution of X_n, given that (X_j, X_n) are records, is*

$$\begin{aligned}
&\Pr(X_n \leq x_2 | R_j = 1, R_n = 1)\\
&= \frac{\Phi_{n-2}(\mathbf{0};\mathbf{\Gamma}_{\mathcal{I}^\complement,\mathcal{I}^\complement})}{\Phi_{n-1}(\mathbf{0};\tilde{\mathbf{\Omega}})} \Psi_{2,n-2}(0, x_2; \boldsymbol{D}\bar{\mathbf{\Sigma}}_{\mathcal{I},\mathcal{I}}\boldsymbol{D}^\top, \mathbf{\Delta}, \bar{\mathbf{\Sigma}}_{\mathcal{I}^\complement,\mathcal{I}^\complement;\mathcal{I}})
\end{aligned}$$

Hence, different to the case of independent random variables in Falk et al. (2018) we have that the distribution of X_n, being a record, is affected by the additional knowledge that at time $j < n$ there was a record.

4.2.3 Asymptotic results for records of stationary sequences

Although stationary Gaussian sequences are useful for a wide range of statistical analysis (e.g., Lindgren 2012, Brockwell and Davis 2013, Banerjee, Carlin, and

Gelfand 2014, Cressie and Wikle 2015, to name a few), a natural question that arises is the following. What are the properties of records for a stationary sequence of dependent rvs when the univariate marginal distribution F, is non-Gaussian? This question is even more relevant if it is assumed that F is unknown, which concerns many real-world applications. Some of the previous results are clearly independent of the underlying df F, provided it is continuous. The probability that X_n is a record, or the distribution of the arrival time of the n-th record, for example, do not depend on F. The distribution of X_n, conditional to the assumption that it is a record, however does depend on F.

Theorem 4.2.11. *Let $\{X_n, n \geq 1\}$ be a stationary sequence that has extremal index $0 < \theta \leq 1$. Then,*

$$n \Pr(R_n = 1) \to \theta^{-1}, \quad \text{as} \quad n \to \infty. \tag{4.19}$$

Furthermore, there are sequences of norming constants $a_n > 0$ and $b_n \in \mathbb{R}$ such that the asymptotic distribution of X_n (suitably normalized), given that it is a record, is

$$\lim_{n\to\infty} \Pr(X_n \leq a_n x + b_n | R_n = 1) = G_\gamma^\theta(x), \quad x, \gamma \in \mathbb{R},\ 0 < \theta \leq 1. \tag{4.20}$$

Proof. First, we show that there are on average approximately θ^{-1} records among $X_1, \ldots, X_n$, for large n. Precisely, (4.19) is obtained from

$$\begin{aligned} n \Pr(X_n > M_{n-1}) &= n \int_{\text{supp}(F)} \Pr(M_{n-1} < v | X_n = v) f_{X_n}(v) \mathrm{d}u \\ &= \int \underbrace{\Pr(M_{n-1} < u_n(t) | X_n = u_n(t))}_{A_n} \times \underbrace{n a_n f_{X_n}(u_n(t)) \mathbb{1}(\mathbb{D})}_{B_n} \mathrm{d}t \end{aligned}$$

where we used the change of variable $v = u_n(t)$ with $u_n(t) = a_n t + b_n$ and $\mathbb{D} := \{t \in \mathbb{R} : u_n(t) \in \text{supp}(F)\}$. By Theorem 3.7.1 in Leadbetter et al. 1983, for any $t \in \mathbb{D}$, we have

$$\begin{aligned} &\Pr(M_{n-1} \leq u_n(t), X_n \leq u_n(t)) \\ &= \Pr(M_{n-1} \leq u_n(t) | X_n \leq u_n(t))(1 - \Pr(X_n > u_n(t))) \\ &= \Pr(M_{n-1} \leq u_n(t) | X_n \leq u_n(t))(1 + o(1)), \end{aligned}$$

and, on the other hand, we have

$$\begin{aligned}\Pr(M_{n-1} \le u_n(t)) &\ge \Pr(M_{n-1} \le u_n(t), X_n \le u_n(t)) \\ &\ge \Pr(M_{n-1} \le u_n(t)) - \Pr(M_{n-1} \le u_n(t), X_n > u_n(t)) \\ &\ge \Pr(M_{n-1} \le u_n(t)) - \Pr(X_n > u_n(t)) \\ &\ge \Pr(M_{n-1} \le u_n(t)) + o(1).\end{aligned}$$

From these two results it follows that for any $t \in \mathbb{D}$ we have

$$\Pr(M_{n-1} \le u_n(t) | X_n \le u_n(t)) = \Pr(M_{n-1} \le u_n(t)) + o(1). \tag{4.21}$$

By (4.21) and Theorem 3.7.1 in Leadbetter et al. 1983 (cf. O'Brien 1987, Rootzén 1988) we have

$$\begin{aligned}A_n &\approx \Pr(M_{n-1} < u_n(t)), \quad n \to \infty \\ &\approx \exp\left\{-n\Pr(X_1 > u_n(t)) \Pr\left(M_{r_n} \le u_n(t) | X_1 > u_n(t)\right)\right\}, \quad \text{as } n \to \infty.\end{aligned}$$

Note that

$$n\Pr(X_1 > u_n(t)) \overset{n\to\infty}{\longrightarrow} V(t) = \begin{cases} t^{-\alpha},\ t > 0,\ \alpha < 0, & \text{if } F \in \mathcal{D}(G_\alpha), \\ (-t)^{\beta},\ t < 0,\ \beta > 0, & \text{if } F \in \mathcal{D}(G_\beta), \\ e^{-t},\ t \in \mathbb{R}, & \text{if } F \in \mathcal{D}(G_0), \end{cases}$$

and

$$\Pr\left(M_{r_n} \le u_n(t)) | X_1 > u_n(t)\right) \overset{n\to\infty}{\longrightarrow} \theta.$$

By Resnick (1987, Ch. 2) we have

$$B_n \overset{n\to\infty}{\longrightarrow} g(t)\mathbb{1}(t \in \text{supp}(G_\gamma)) = \begin{cases} \alpha t^{-(\alpha+1)},\ t > 0,\ \alpha < 0, & \text{if } F \in \mathcal{D}(G_\alpha), \\ \beta(-t)^{\beta-1},\ t < 0,\ \beta > 0, & \text{if } F \in \mathcal{D}(G_\beta), \\ e^{-t},\ t \in \mathbb{R}, & \text{if } F \in \mathcal{D}(G_0). \end{cases}$$

Therefore, putting all these results together we obtain, as $n \to \infty$,

$$n\Pr(X_n > M_{n-1}) \approx \int_{\text{supp}(G_\gamma)} \exp\left\{-V(t)\theta\right\} g(t) \mathrm{d}t = \theta^{-1},$$

and hence, (4.19) is proven.

Finally, using similar arguments we obtain

$$\begin{aligned}
&\Pr(X_n \le a_n x + b_n | R_n = 1)\\
&= \frac{\Pr(X_n \le a_n x + b_n, X_n > M_{n-1})}{\Pr(X_n > M_{n-1})}\\
&= \frac{n \int_{v \in \mathrm{supp}(F): v \le a_n x + b_n} \Pr(M_{n-1} < v | X_n = v) f_{X_n}(v) \mathrm{d}u}{n \Pr(X_n > M_{n-1})}\\
&\overset{n\to\infty}{\longrightarrow} G_\gamma^\theta(x), \quad x, \gamma \in \mathbb{R}
\end{aligned}$$

and the proof is complete. □

Theorem 4.2.11 states that for a stationary sequence of dependent rvs $\{X_n, n \ge 1\}$, under appropriate conditions on the dependence structure, the asymptotic distribution of X_n (appropriately normalized), being a record, coincides with the asymptotic distribution G_γ^θ of the normalized maximum. This finding generalizes Lemma 2.1 in Falk et al. (2018), derived for a sequence of indepedent rvs. Indeed, the same result is obtained for $\theta = 1$.

In the following part the are three specific examples of asymptotic distributions of records that stem from the general formula (4.20) in Theorem 4.2.11.

Example 4.2.12 (Chernick et al. 1981). *For an integer $m \ge 2$, let $\{\varepsilon_n, n \ge 1\}$ be a sequence of i.i.d. rvs uniformly distributed on $\{0, 1/m, \dots, (m-1)/m\}$. Let X_0 be a rv uniformly distributed on $[0, 1]$, being independent of $\{\varepsilon_n\}$. The process*

$$X_n = m^{-1} X_{n-1} + \varepsilon_n, \quad n \ge 1,$$

defines a strictly stationary first-order autoregressive sequence. For $n = 1, 2, \dots$ take the norming constants $a_n > 1/n$ and $b_n = 1$. Then,

$$\lim_{n\to\infty} \Pr(X_n \le 1 + x/n | R_n = 1) = e^{\theta x}, \quad x < 0,$$

where $\theta = (m-1)/m$ with $m \ge 2$.

Example 4.2.13 (Hsing et al. 1996). *Let $\{X_{n,i}, n \ge 1, i \ge 0\}$ be a triangular array of rvs such that for every n $\{X_{n,i}, i \ge 0\}$ is a SSG sequence. Define $\rho_{n,j} = \mathrm{E}(X_{n,i} X_{n,i+j})$ with $i \le n$ and $j \ge 1$. Assume that $(1 - \rho_{n,j}) \log n \to \delta_j \in (0, \infty]$*

for all $j \geq 1$ *as* $n \to \infty$. *For* $n = 1, 2, \ldots$ *choose the norming constants* $a_n = (2 \log n)^{-1/2}$ *and*

$$b_n = a_n^{-1} - a_n(\log \log n + \log 4\pi)/2$$

Then,

$$\lim_{n\to\infty} \Pr(X_n \leq a_n x + b_n | R_n = 1) = e^{-\theta e^{-x}}, \quad x \in \mathbb{R},$$

where

$$\theta = \mathrm{E}_U \left\{ \Phi_{|K|} \left(\sqrt{\delta_k} - \frac{U}{2\sqrt{\delta_k}}; \mathbf{\Sigma} \right) \right\}$$

and where $K = \{k \in A \subset \{1, 2 \ldots\} : \delta_k < \infty\}$, U *is a standard exponential rv and* $\mathbf{\Sigma}$ *is a correlation matrix with upper diagonal entries*

$$\frac{\delta_i + \delta_j - \delta_{|i-j|}}{2\sqrt{\delta_i \delta_j}}, \quad 1 \leq i < j \leq |K|.$$

Example 4.2.14 (Leadbetter et al. 1983 Ch. 3.8). *Let* $\varepsilon_1, \varepsilon_2, \ldots$ *be i.i.d. stable* $(1, \alpha, \kappa)$ *rvs. We recall that a rv is stable* (τ, α, κ) *with* $\tau \leq 0$, $0 < \alpha \leq 2$ *and* $|\kappa| \leq 1$ *if its characteristic function is*

$$\omega(x) = \exp\left\{ -\tau^\alpha |x|^\alpha \left(1 - i\frac{\kappa h(x, \alpha) x}{|x|} \right) \right\}$$

where $i^2 = -1$ *and* $h(x, \alpha) = \tan(\pi\alpha/2)$ *for* $\alpha \neq 1$ *and* $h(x, 1) = 2\pi^{-1} \log |x|$ *otherwise. Let* $\{c_i, i \in \mathbb{Z}\}$ *be a sequence of constants satisfying* $\sum_{i=-\infty}^{\infty} |c_i|^\alpha < \infty$ *and* $\sum_{i=-\infty}^{\infty} c_i \log |c_i|$ *is convergent for* $\alpha = 1$ *and* $\kappa \neq 0$. *Define the moving average process*

$$X_n = \sum_{i=-\infty}^{\infty} c_i \varepsilon_{n-i}, \quad n \geq 1.$$

For $n = 1, 2, \ldots$ *choose the norming constants* $a_n = n^{1/\alpha}$ *and* $b_n = 0$. *Then,*

$$\lim_{n\to\infty} \Pr(X_n \leq x n^{1/\alpha} | R_n = 1) = e^{-\theta x^{-\alpha}}, \quad x > 0,$$

where

$$\theta = k_\alpha (c_+^\alpha (1 + \kappa) + c_-^\alpha (1 - \kappa)),$$

with $c_\pm = \max_{-\infty < i < \infty} c_i^\pm$, $c^\pm = \max(0, \pm c_i)$ *and* $k_\alpha = \pi^{-1}\Gamma(\alpha) \sin(\alpha\pi/2)$.

What is the expected number of records that will take place in the case of a stationary sequence of rvs that have extremal index $0 < \theta \leq 1$? We know that

$$\Pr(X_n > M_{n-1}) \approx \frac{1}{n\theta}, \quad n \to \infty.$$

Therefore, by elementary arguments,

$$\mathrm{E}(N) = \sum_{n=1}^{\infty} \Pr(X_n > M_{n-1}) \overset{n\to\infty}{\longrightarrow} \infty.$$

4.3 Records of dependent multivariate Gaussian sequences

Let $\{\boldsymbol{X}_n, n \geq 1\}$ be a sequence of d-dimensional random vectors $\boldsymbol{X}_n = (X_n^{(i)})_{i=1,\ldots,d} \in \mathbb{R}^d$. Here we consider a second-order stationary multivariate Gaussian process and we extend some of the results derived in Section 4.2.2 to the multivariate case. Precisely, we study the probability that a complete record $\boldsymbol{X}_n$ occurs and the distribution of $\boldsymbol{X}_n$ (being a record). We also study the probability that two complete records $(\boldsymbol{X}_j, \boldsymbol{X}_n)$ occur and the joint distribution of $(\boldsymbol{X}_j, \boldsymbol{X}_n)$ (being records). Without loss of generality, assume for simplicity that $\mathrm{E}(\boldsymbol{X}_i) = 0$, $\mathrm{E}(\boldsymbol{X}_i^2) = 1$ for every $1 \leq i \leq n$.

Let $\boldsymbol{X} = (\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n)$ be an nd-dimensional random vector and consider the partition $\boldsymbol{X} = (\boldsymbol{X}_{\mathcal{I}}^\top, \boldsymbol{X}_{\mathcal{I}^\complement}^\top)^\top \sim N_{nd}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with corresponding partition of the parameters $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$. The formula of the conditional distribution of $\boldsymbol{X}_{\mathcal{I}^\complement}$ given that $\boldsymbol{X}_{\mathcal{I}} = \boldsymbol{x}_{\mathcal{I}}$, for all $\boldsymbol{x}_{\mathcal{I}} \in \mathbb{R}^{|\mathcal{I}|}$, in (4.5) is still valid with the obvious changes. Further on we will provide the specific details whenever we use such a formula.

Proposition 4.3.1. *Let $\{\boldsymbol{X}_n, n \geq 1\}$ be a SSG sequence of random vectors in $\mathbb{R}^d$. For every $n \geq 2$, the probability that $\boldsymbol{X}_n$ is a record and the distribution of $\boldsymbol{X}_n$, given that it is a record, are equal to*

$$\Pr(R_n^{CR} = 1) = \Phi_{(n-1)d}(\boldsymbol{0}; \boldsymbol{\Gamma}_{1:n-1,n}), \tag{4.22}$$

$$\Pr(\boldsymbol{X}_n \leq \boldsymbol{x} | R_n^{CR} = 1) = \Psi_{d,(n-1)d}\left(\boldsymbol{x}; \bar{\boldsymbol{\Sigma}}_n, \boldsymbol{\varrho}_{1:n-1,n}, \bar{\boldsymbol{\Sigma}}_{1:n-1,1:n-1;n}\right), \tag{4.23}$$

where

$$\boldsymbol{B} = \begin{pmatrix} \mathbf{1}_{n-1} & \mathbf{0}_{n-1} & \dots & \mathbf{0}_{n-1} \\ \mathbf{0}_{n-1} & \mathbf{1}_{n-1} & \dots & \mathbf{0}_{n-1} \\ \vdots & \vdots & & \vdots \\ \mathbf{0}_{n-1} & \mathbf{0}_{n-1} & \dots & \mathbf{1}_{n-1} \end{pmatrix} \in \mathbb{R}^{(n-1)d,d}$$

$$\boldsymbol{\varrho}_{1:n-1,n} = \boldsymbol{\sigma}^{-1}_{1:n-1,1:n-1;n}(\boldsymbol{B} - \boldsymbol{\Sigma}^{(1:d)}_{1:n-1,n}\bar{\boldsymbol{\Sigma}}_n^{-1}) \in \mathbb{R}^{(n-1)d,d} \tag{4.24}$$

where $\boldsymbol{\Sigma}^{(1:d)}_{1:n-1,n}$ *is the covariance matrix of*

$$(X_1^{(1)}, X_2^{(1)}, \dots, X_{n-1}^{(1)}, \dots, X_1^{(d)}, X_2^{(d)}, \dots, X_{n-1}^{(d)})$$

and $\boldsymbol{X}_n$*,* $\bar{\boldsymbol{\Sigma}}_n$ *is the variance-covariance matrix of* $\boldsymbol{X}_n$ *and*

$$\boldsymbol{\Gamma}_{1:n-1,n} = \bar{\boldsymbol{\Sigma}}_{1:n-1,1:n-1;n} + \boldsymbol{\varrho}_{1:n-1,n}\bar{\boldsymbol{\Sigma}}_n\boldsymbol{\varrho}^\top_{1:n-1,n}$$

Proof. We start deriving the probability that $\boldsymbol{X}_n$ is a record.

$$\begin{aligned}\Pr(\boldsymbol{X}_n > \boldsymbol{M}_{n-1}) &= \Pr(X_{n,i} > M_{n-1,i}, i = 1, \dots, d) \\ &= \int_{\mathbb{R}^d} \Pr(M_{n-1,i} < z_i, i = 1, \dots, d | X_{n,i} = z_i, i = 1, \dots, d)\phi_d(\boldsymbol{z})\mathrm{d}\boldsymbol{z}\end{aligned}$$

Let $\boldsymbol{X}^{(i)}_{1:n-1}$ be the vector of the i-th components of $\boldsymbol{X}_1, \dots, \boldsymbol{X}_{n-1}$. Then,

$$\begin{aligned}&\Pr(M_{n-1,i} < z_i, i = 1, \dots, d | X_{n,i} = z_i, i = 1, \dots, d) \\ &= \Pr(\cap_{i=1}^d \{X_{k,i} < z_i, k = 1, \dots, n-1\} | X_{n,i} = z_i, i = 1, \dots, d) \\ &= \Pr(\boldsymbol{X}^{(i)}_{1:n-1} < \mathbf{1}_{n-1} z_i, i = 1, \dots, d | \boldsymbol{X}_n = \boldsymbol{z}).\end{aligned}$$

By the multivariate version of the conditional Gaussian distribution in (4.5) we have

$$(\boldsymbol{X}_1, \boldsymbol{X}_2, \dots, \boldsymbol{X}_{n-1} | \boldsymbol{X}_n = \boldsymbol{z}_n) \sim N_{(n-1)d}\left(\boldsymbol{\mu}_{1:n-1,n}, \boldsymbol{\Sigma}_{1:n-1,1:n-1;n}\right)$$

where

$$\begin{aligned}\boldsymbol{\mu}_{1:n-1,n} &= \left(\boldsymbol{\Sigma}^{(i)}_{1:n-1,n}\bar{\boldsymbol{\Sigma}}_n^{-1}\right)_{i=1,\dots,d}\boldsymbol{z} \in \mathbb{R}^{(n-1)d} \\ \boldsymbol{\Sigma}_{1:n-1,1:n-1;n} &= \left(\boldsymbol{\Sigma}^{(i,h)}_{1:n-1,1:n-1} - \boldsymbol{\Sigma}^{(i)}_{1:n-1,n}\bar{\boldsymbol{\Sigma}}_n^{-1}{\boldsymbol{\Sigma}^{(h)}}^\top_{1:n-1,n}\right)_{i,h=1,\dots,d}.\end{aligned}$$

$\boldsymbol{\Sigma}^{(i)}_{1:n-1,n}$ is the covariance matrix of $\boldsymbol{X}^{(i)}_{1:n-1}$ and $\boldsymbol{X}_n$, and $\boldsymbol{\Sigma}^{(i,h)}_{1:n-1,1:n-1}$ is the covariance matrix of $\boldsymbol{X}^{(i)}_{1:n-1}$ and $\boldsymbol{X}^{(h)}_{1:n-1}$. We have that

$$\Pr(\boldsymbol{X}^{(i)}_{1:n-1} < \mathbf{1}_{n-1}z_i, i = 1, \dots, d | \boldsymbol{X}_n = \boldsymbol{z}) = \Phi_{(n-1)d}(\boldsymbol{\varrho}_{1:n-1,n}\boldsymbol{z})$$

where $\boldsymbol{\varrho}_{1:n-1,n}$ is defined as in equation (4.24). Therefore

$$\begin{aligned}\Pr(\boldsymbol{X}_n > \boldsymbol{M}_{n-1}) &= \int_{\mathbb{R}^d} \Phi_{(n-1)d}(\boldsymbol{\varrho}_{1:n-1,n}\boldsymbol{z})\phi_d(\boldsymbol{z})\mathrm{d}\boldsymbol{z} \\ &= \mathrm{E}\left(\Phi_{(n-1)d}(\boldsymbol{\varrho}_{1:n-1,n}\boldsymbol{Z})\right),\end{aligned}$$

where $\boldsymbol{Z} \sim N_{(n-1)d}(\mathbf{0}, \bar{\boldsymbol{\Sigma}}_{1:n-1,1:n-1;n})$ and the claim follows by applying Proposition 7.1 in Azzalini and Capitanio (1999).

The computation of the distribution function follows the same procedure. We need to compute

$$\begin{aligned}&\Pr(\boldsymbol{X}_n \leq \boldsymbol{x}, \boldsymbol{X}_n > \boldsymbol{M}_{n-1}) \\ &= \int_{(-\infty,\boldsymbol{x}]} \Pr(M_{n-1,i} < z_i, i = 1, \dots, d | X_{n,i} = z_i, i = 1, \dots, d)\phi_d(\boldsymbol{z})\mathrm{d}\boldsymbol{z} \\ &= \Phi_{(n-1)d}(\mathbf{0}; \bar{\boldsymbol{\Sigma}}_{1:n-1;n} + \boldsymbol{\varrho}_{1:n-1,n}\bar{\boldsymbol{\Sigma}}_n\boldsymbol{\varrho}^\top_{1:n-1,n}) \\ &\quad \times \Psi_{d,(n-1)d}(\boldsymbol{x}; \bar{\boldsymbol{\Sigma}}_n, \boldsymbol{\varrho}_{1:n-1,n}, \bar{\boldsymbol{\Sigma}}_{1:n-1;n})\end{aligned}$$

□

REMARK 4.3.2. *For every* $n = 1, 2 \dots$, $m < n$ *and* $1 \leq i < j \leq d$, *if* $\mathrm{Cor}(X^{(i)}_n, X^{(j)}_n) = 0$ *and* $\mathrm{Cor}(X^{(i)}_n, X^{(i)}_m) = 0$, *then we obtain*

$$\Pr\left(\boldsymbol{X}_n > \boldsymbol{M}_{n-1}\right) = \Phi_{(n-1)d}(\mathbf{0}; \bar{\boldsymbol{\Sigma}}_{1:n-1,1:n-1;n} + \boldsymbol{\varrho}_{1:n-1,n}\boldsymbol{\varrho}^\top_{1:n-1,n})$$

where the variance-covariance matrix is a diagonal block matrix, with each diagonal block being equal to $\boldsymbol{I}_{n-1} + \mathbf{1}_{n-1}\mathbf{1}^\top_{n-1}$. *Since we have* d *diagonal blocks, we obtain*

$$\begin{aligned}\Phi_{(n-1)d}(\mathbf{0}; \bar{\boldsymbol{\Sigma}}_{1:n-1,1:n-1;n} + \boldsymbol{\varrho}_{1:n-1,n}\boldsymbol{\varrho}^\top_{1:n-1,n}) &= \left(\Phi_{n-1}(\mathbf{0}; \boldsymbol{I}_{n-1} + \mathbf{1}_{n-1}\mathbf{1}^\top_{n-1})\right)^d \\ &= n^{-d}\end{aligned}$$

by the result in Remark 4.2.2.

Our next result deals with the joint distribution of two complete records at times j and $n > j$. In the following, we use the notation $\boldsymbol{x}_{\mathcal{J}}$, $\mathcal{J} \subseteq \{1, \ldots, d\}$ to indicate a vector of dimension $|\mathcal{J}|$ whose components are the entries of $\boldsymbol{x} \in \mathbb{R}^d$ determined by the elements of $\mathcal{J}$.

Theorem 4.3.3. *Let $\{\boldsymbol{X}_n, n \geq 1\}$ be a SSG sequence of random vectors in $\mathbb{R}^d$. For j and $n > j$, set $\mathcal{I} = \{j, n\}$. Then*

$$\Pr(R_j^{CR} = 1, R_n^{CR} = 1) = \Phi_{d(n-2)}(\mathbf{0}; \boldsymbol{\Gamma}_{\mathcal{I}^\complement \mathcal{I}^\complement})\Psi_{d,d(n-2)}(\mathbf{0}; \mathcal{L}_1) \tag{4.25}$$

$$\begin{aligned} &\Pr(\boldsymbol{X}_j \leq \boldsymbol{x}_1,\, \boldsymbol{X}_n \leq \boldsymbol{x}_2 | R_j^{CR} = 1, R_n^{CR} = 1) \\ &= \sum_{\mathcal{J} \subseteq \{1,\ldots,d\}} \Psi_{d,d(n-2)}(\mathbf{0}; \mathcal{L}_1)^{-1} \big(\Psi_{2d,d(n-2)}(\mathbf{0}_{\mathcal{J}}, \boldsymbol{x}_{1\bar{\mathcal{J}}}, \boldsymbol{x}_{1\mathcal{J}}, \boldsymbol{x}_{2\bar{\mathcal{J}}}; \mathcal{L}_{\mathcal{J}}) \\ &\qquad - \Psi_{2d,d(n-2)}(\mathbf{0}_{\mathcal{J}}, \boldsymbol{x}_{1\bar{\mathcal{J}}}, \boldsymbol{x}_{1\mathcal{J}}, \boldsymbol{x}_{1\bar{\mathcal{J}}}; \mathcal{L}_{\mathcal{J}})\big), \end{aligned} \tag{4.26}$$

where $\Psi_{m,q}(\cdot; \mathcal{L}) \sim CSN_{m,q}(\mathcal{L})$, $\boldsymbol{\Gamma}_{\mathcal{I}^\complement \mathcal{I}^\complement} := \bar{\boldsymbol{\Sigma}}_{\mathcal{I}^\complement \mathcal{I}^\complement; \mathcal{I}} + \boldsymbol{\varrho}_{1:n-1,n} \bar{\boldsymbol{\Sigma}}_{\mathcal{I}\mathcal{I}} \boldsymbol{\varrho}_{1:n-1,n}^\top$,

$$\mathcal{L}_{(\cdot)} = \left(\boldsymbol{D}_{(\cdot)} \bar{\boldsymbol{\Sigma}}_{\mathcal{I}\mathcal{I}} \boldsymbol{D}_{(\cdot)}^\top, \boldsymbol{\Delta}_{(\cdot)}, \bar{\boldsymbol{\Sigma}}_{\mathcal{I}^\complement \mathcal{I}^\complement; \mathcal{I}}\right), \tag{4.27}$$

$\boldsymbol{\Delta}_{(\cdot)} = \boldsymbol{\varrho}_{1:n-1,n} \bar{\boldsymbol{\Sigma}}_{\mathcal{I}^\complement, \mathcal{I}^\complement} \boldsymbol{D}_{(\cdot)}^\top (\boldsymbol{D}_{(\cdot)} \bar{\boldsymbol{\Sigma}}_{\mathcal{I},\mathcal{I}} \boldsymbol{D}_{(\cdot)}^\top)^{-1}$ *and*

$$D_1 = \begin{pmatrix} \boldsymbol{I}_d & -\boldsymbol{I}_d \end{pmatrix} \qquad D_{\mathcal{J}} = \begin{pmatrix} \boldsymbol{I}_{\mathcal{J}} & \mathbf{0}_{\bar{\mathcal{J}}} & -\boldsymbol{I}_{\mathcal{J}} & \mathbf{0}_{\bar{\mathcal{J}}} \\ \mathbf{0}_{\mathcal{J}} & \boldsymbol{I}_{\bar{\mathcal{J}}} & \mathbf{0}_{\mathcal{J}} & \mathbf{0}_{\bar{\mathcal{J}}} \\ \mathbf{0}_{\mathcal{J}} & \mathbf{0}_{\bar{\mathcal{J}}} & \boldsymbol{I}_{\mathcal{J}} & \mathbf{0}_{\bar{\mathcal{J}}} \\ \mathbf{0}_{\mathcal{J}} & \mathbf{0}_{\bar{\mathcal{J}}} & \mathbf{0}_{\mathcal{J}} & \boldsymbol{I}_{\bar{\mathcal{J}}} \end{pmatrix}$$

Proof. We compute

$$\begin{aligned} &\Pr(R_j^{\text{CR}} = 1, R_n^{\text{CR}} = 1) \\ &= \Pr(\boldsymbol{X}_j > \boldsymbol{M}_{j-1}, \boldsymbol{X}_n > \boldsymbol{M}_{n-1}) \\ &= \Pr(\boldsymbol{X}_j > \boldsymbol{M}_{j-1}, \boldsymbol{X}_n > \boldsymbol{M}_{j+1:n-1}, \boldsymbol{X}_j < \boldsymbol{X}_n) \\ &= \int_{\mathbb{R}^d} \int_{(-\infty, \boldsymbol{z}_n]} \Pr(\boldsymbol{M}_{j-1} < \boldsymbol{z}_j, \boldsymbol{M}_{j+1:n-1} < \boldsymbol{z}_n | \boldsymbol{X}_i = \boldsymbol{z}_i, i \in \mathcal{I}) \phi_{2d}(\boldsymbol{z}_j, \boldsymbol{z}_n; \bar{\boldsymbol{\Sigma}}_{\mathcal{I}\mathcal{I}}) \mathrm{d}\boldsymbol{z}_j \mathrm{d}\boldsymbol{z}_n \end{aligned}$$

First of all, we recall the inverse blok-matrix of a two-by-two block matrix:

$$\bar{\boldsymbol{\Sigma}}_{\mathcal{I}\mathcal{I}}^{-1} = \begin{pmatrix} \boldsymbol{\Lambda}_1 & -\bar{\boldsymbol{\Sigma}}_j^{-1} \boldsymbol{\Sigma}_{j,n} \boldsymbol{\Lambda}_2 \\ -\boldsymbol{\Lambda}_2 \boldsymbol{\Sigma}_{n,j} \bar{\boldsymbol{\Sigma}}_j^{-1} & \boldsymbol{\Lambda}_2 \end{pmatrix}$$

where $\boldsymbol{\Lambda}_i$ is the Schur complement of $\bar{\boldsymbol{\Sigma}}_i$ in $\bar{\boldsymbol{\Sigma}}_{\mathcal{I}\mathcal{I}}$, for $i \in \mathcal{I}$, for example, $\boldsymbol{\Lambda}_1 = \left(\bar{\boldsymbol{\Sigma}}_j - \boldsymbol{\Sigma}_{j,n}\bar{\boldsymbol{\Sigma}}_n^{-1}\boldsymbol{\Sigma}_{n,j}\right)^{-1}$. By the multivariate version of the conditional Gaussian df in (4.5) we have that $(\boldsymbol{X}_i, i \in \mathcal{I}^{\complement})|(\boldsymbol{X}_i = \boldsymbol{z}_i, i \in \mathcal{I}) \sim N_{|\mathcal{I}^{\complement}|}\left(\boldsymbol{\mu}_{\mathcal{I}^{\complement}\mathcal{I}^{\complement};\mathcal{I}}, \boldsymbol{\Sigma}_{\mathcal{I}^{\complement}\mathcal{I}^{\complement};\mathcal{I}}\right)$. Specifically we have $\boldsymbol{\mu}_{\mathcal{I}^{\complement}\mathcal{I}^{\complement};\mathcal{I}} = \left((\boldsymbol{\mu}_i^\top, i = 1, \ldots, d), (\boldsymbol{\mu}_h^\top, h = d+1, \ldots, 2d)\right)^\top$ with

$$\begin{aligned}
\boldsymbol{\mu}_i &= \left(\boldsymbol{\Sigma}^{(i)}_{1:j-1,j}\boldsymbol{\Lambda}_1 - \boldsymbol{\Sigma}^{(i)}_{1:j-1,n}\boldsymbol{\Lambda}_2\boldsymbol{\Sigma}_{n,j}\bar{\boldsymbol{\Sigma}}_j^{-1}\right)\boldsymbol{z}_j \\
&\quad + \left(-\boldsymbol{\Sigma}^{(i)}_{1:j-1,j}\bar{\boldsymbol{\Sigma}}_j^{-1}\boldsymbol{\Sigma}_{j,n}\boldsymbol{\Lambda}_2 + \boldsymbol{\Sigma}^{(i)}_{1:j-1,n}\boldsymbol{\Lambda}_2\right)\boldsymbol{z}_n \\
&=: \boldsymbol{\mu}_{ij}\boldsymbol{z}_j + \boldsymbol{\mu}_{in}\boldsymbol{z}_n,
\end{aligned} \tag{4.28}$$

$$\begin{aligned}
\boldsymbol{\mu}_h &= \left(\boldsymbol{\Sigma}^{(h-d)}_{j+1:n-1,j}\boldsymbol{\Lambda}_1 - \boldsymbol{\Sigma}^{(h-d)}_{j+1:n-1,n}\boldsymbol{\Lambda}_2\boldsymbol{\Sigma}_{n,j}\bar{\boldsymbol{\Sigma}}_j^{-1}\right)\boldsymbol{z}_j \\
&\quad + \left(-\boldsymbol{\Sigma}^{(h-d)}_{j+1:n-1,j}\bar{\boldsymbol{\Sigma}}_j^{-1}\boldsymbol{\Sigma}_{j,n}\boldsymbol{\Lambda}_2 + \boldsymbol{\Sigma}^{(h-d)}_{j+1:n-1,n}\boldsymbol{\Lambda}_2\right)\boldsymbol{z}_n \\
&=: \boldsymbol{\mu}_{hj}\boldsymbol{z}_j + \boldsymbol{\mu}_{hn}\boldsymbol{z}_n,
\end{aligned} \tag{4.29}$$

and $\boldsymbol{\Sigma}_{\mathcal{I}^{\complement}\mathcal{I}^{\complement};\mathcal{I}}$ is $(2d) \times (2d)$ matrix. It is defined by blocks $(\boldsymbol{\Sigma}_{\mathcal{I}^{\complement}\mathcal{I}^{\complement};\mathcal{I}})_{i,h=1,\ldots,2d}$ of the form

$$\begin{aligned}
&(\boldsymbol{\Sigma}_{\mathcal{I}^{\complement}\mathcal{I}^{\complement};\mathcal{I}})_{ih} \qquad \text{for } i, h = 1, \ldots, d \\
&= \boldsymbol{\Sigma}^{(i,h)}_{1:j-1,1:j-1} - \left(\left(\boldsymbol{\Sigma}^{(i)}_{1:j-1,j}\boldsymbol{\Lambda}_1 - \boldsymbol{\Sigma}^{(i)}_{1:j-1,n}\boldsymbol{\Lambda}_2\boldsymbol{\Sigma}_{n,j}\bar{\boldsymbol{\Sigma}}_j^{-1}\right)\boldsymbol{\Sigma}^{(h)}_{1:j-1,j} \right.\\
&\quad \left. + \left(-\boldsymbol{\Sigma}^{(i)}_{1:j-1,j}\bar{\boldsymbol{\Sigma}}_j^{-1}\boldsymbol{\Sigma}_{j,n}\boldsymbol{\Lambda}_2 + \boldsymbol{\Sigma}^{(i)}_{1:j-1,n}\boldsymbol{\Lambda}_2\right)\boldsymbol{\Sigma}^{(h)}_{1:j-1,n}\right)_{ih}
\end{aligned}$$

$$\begin{aligned}
&(\boldsymbol{\Sigma}_{\mathcal{I}^{\complement}\mathcal{I}^{\complement};\mathcal{I}})_{ih} \qquad \text{for } i = 1, \ldots, d,\ h = d+1, \ldots, 2d \\
&= \boldsymbol{\Sigma}^{(i,h-d)}_{1:j-1,j+1:n-1} - \left(\left(\boldsymbol{\Sigma}^{(i)}_{1:j-1,j}\boldsymbol{\Lambda}_1 - \boldsymbol{\Sigma}^{(i)}_{1:j-1,n}\boldsymbol{\Lambda}_2\boldsymbol{\Sigma}_{n,j}\bar{\boldsymbol{\Sigma}}_j^{-1}\right)\boldsymbol{\Sigma}^{(h-d)}_{j+1:n-1,j} \right.\\
&\quad \left. + \left(-\boldsymbol{\Sigma}^{(i)}_{1:j-1,j}\bar{\boldsymbol{\Sigma}}_j^{-1}\boldsymbol{\Sigma}_{j,n}\boldsymbol{\Lambda}_2 + \boldsymbol{\Sigma}^{(i)}_{1:j-1,n}\boldsymbol{\Lambda}_2\right)\boldsymbol{\Sigma}^{(h-d)}_{j+1:n-1,n}\right)_{ih}
\end{aligned}$$

$$\begin{aligned}
&(\boldsymbol{\Sigma}_{\mathcal{I}^{\complement}\mathcal{I}^{\complement};\mathcal{I}})_{ih} \qquad \text{for } i = d+1, \ldots, 2d,\ h = 1, \ldots, d \\
&= \boldsymbol{\Sigma}^{(i-d,h)}_{j+1:n-1,1:j-1} - \left(\left(\boldsymbol{\Sigma}^{(i-d)}_{j+1:n-1,j}\boldsymbol{\Lambda}_1 - \boldsymbol{\Sigma}^{(i-d)}_{j+1:n-1,n}\boldsymbol{\Lambda}_2\boldsymbol{\Sigma}_{n,j}\bar{\boldsymbol{\Sigma}}_j^{-1}\right)\boldsymbol{\Sigma}^{(h)}_{1:j-1,j} \right.\\
&\quad \left. + \left(-\boldsymbol{\Sigma}^{(i-d)}_{j+1:n-1,j}\bar{\boldsymbol{\Sigma}}_j^{-1}\boldsymbol{\Sigma}_{j,n}\boldsymbol{\Lambda}_2 + \boldsymbol{\Sigma}^{(i-d)}_{j+1:n-1,j}\boldsymbol{\Lambda}_2\right)\boldsymbol{\Sigma}^{(h)}_{1:j-1,n}\right)_{ih}
\end{aligned}$$

$$\begin{aligned}
&(\boldsymbol{\Sigma}_{\mathcal{I}^{\complement}\mathcal{I}^{\complement};\mathcal{I}})_{ih} \qquad \text{for } i, h = d+1, \ldots, 2d \\
&= \boldsymbol{\Sigma}^{(i-d,h-d)}_{j+1:n-1,j+1:n-1} - \left(\left(\boldsymbol{\Sigma}^{(i-d)}_{j+1:n-1,j}\boldsymbol{\Lambda}_1 - \boldsymbol{\Sigma}^{(i-d)}_{j+1:n-1,n}\boldsymbol{\Lambda}_2\boldsymbol{\Sigma}_{n,j}\bar{\boldsymbol{\Sigma}}_j^{-1}\right)\boldsymbol{\Sigma}^{(h-d)}_{j+1:n-1,j} \right.\\
&\quad \left. + \left(-\boldsymbol{\Sigma}^{(i-d)}_{j+1:n-1,j}\bar{\boldsymbol{\Sigma}}_j^{-1}\boldsymbol{\Sigma}_{j,n}\boldsymbol{\Lambda}_2 + \boldsymbol{\Sigma}^{(i-d)}_{j+1:n-1,j}\boldsymbol{\Lambda}_2\right)\boldsymbol{\Sigma}^{(h-d)}_{j+1:n-1,n}\right)_{ih}
\end{aligned}$$

Therefore, we obtain

$$\begin{aligned}&\Pr(\boldsymbol{X}^{(i)}_{1:j-1} < \mathbf{1}_{j-1}\boldsymbol{z}_j, \boldsymbol{X}^{(i)}_{j+1:n-1} < \mathbf{1}_{n-j-1}\boldsymbol{z}_n, i = 1, \dots, d | \boldsymbol{X}_i = \boldsymbol{z}_i, i \in \mathcal{I})\\ &= \Pr(\boldsymbol{Z} < \boldsymbol{\varrho}_{\mathcal{I}^\complement,\mathcal{I}^\complement}\boldsymbol{z})\end{aligned}$$

where $\boldsymbol{Z} \sim N_{(n-2)d}(\mathbf{0}, \bar{\boldsymbol{\Sigma}}_{\mathcal{I}^\complement\mathcal{I}^\complement;\mathcal{I}})$,

$$\boldsymbol{\varrho}_{\mathcal{I}^\complement,\mathcal{I}^\complement} = \boldsymbol{\sigma}^{-1}_{\mathcal{I}^\complement,\mathcal{I}^\complement;\mathcal{I}}(\boldsymbol{B} - \boldsymbol{\Sigma}^{(1:d)}_{\mathcal{I}^\complement,\mathcal{I}}\bar{\boldsymbol{\Sigma}}^{-1}_{\mathcal{I}\mathcal{I}}) \in \mathbb{R}^{(n-2)d,d} \tag{4.30}$$

$$\boldsymbol{B} = \begin{pmatrix} \mathbf{1}_{j-1} & \mathbf{0}_{j-1} & \dots & \mathbf{0}_{j-1} & \mathbf{0}_{j-1} & \mathbf{0}_{j-1} & \dots & \mathbf{0}_{j-1} \\ \mathbf{0}_{j-1} & \mathbf{1}_{j-1} & \dots & \mathbf{0}_{j-1} & \mathbf{0}_{j-1} & \mathbf{0}_{j-1} & \dots & \mathbf{0}_{j-1} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \mathbf{0}_{j-1} & \mathbf{0}_{j-1} & \dots & \mathbf{1}_{j-1} & \mathbf{0}_{j-1} & \mathbf{0}_{j-1} & \dots & \mathbf{0}_{j-1} \\ \mathbf{0}_{n-j-1} & \mathbf{0}_{n-j-1} & \dots & \mathbf{0}_{n-j-1} & \mathbf{1}_{n-j-1} & \mathbf{0}_{n-j-1} & \dots & \mathbf{0}_{n-j-1} \\ \mathbf{0}_{n-j-1} & \mathbf{0}_{n-j-1} & \dots & \mathbf{0}_{n-j-1} & \mathbf{0}_{n-j-1} & \mathbf{1}_{n-j-1} & \dots & \mathbf{0}_{n-j-1} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \mathbf{0}_{n-j-1} & \mathbf{0}_{n-j-1} & \dots & \mathbf{0}_{n-j-1} & \mathbf{0}_{n-j-1} & \mathbf{0}_{n-j-1} & \dots & \mathbf{1}_{n-j-1} \end{pmatrix}$$

and $\boldsymbol{z} = (\boldsymbol{z}_j, \boldsymbol{z}_n)$ is a column vector with length $2d$. We obtain

$$\begin{aligned}&\Pr(R^{\mathrm{CR}}_j = 1, R^{\mathrm{CR}}_n = 1)\\ &= \int_{\mathbb{R}^d}\int_{(-\boldsymbol{\infty},\boldsymbol{z}_n]} \Phi_{d(n-2)}(\boldsymbol{\varrho}_{\mathcal{I}^\complement,\mathcal{I}^\complement}\boldsymbol{z}; \bar{\boldsymbol{\Sigma}}_{\mathcal{I}^\complement\mathcal{I}^\complement;\mathcal{I}})\phi_{2d}(\boldsymbol{z}_j, \boldsymbol{z}_n; \bar{\boldsymbol{\Sigma}}_{\mathcal{I}\mathcal{I}})\mathrm{d}\boldsymbol{z}_j\mathrm{d}\boldsymbol{z}_n\\ &= \Phi_{d(n-2)}(\mathbf{0}; \bar{\boldsymbol{\Sigma}}_{\mathcal{I}^\complement\mathcal{I}^\complement;\mathcal{I}} + \boldsymbol{\varrho}_{\mathcal{I}^\complement,\mathcal{I}^\complement}\bar{\boldsymbol{\Sigma}}_{\mathcal{I}\mathcal{I}}\boldsymbol{\varrho}^\top_{\mathcal{I}^\complement,\mathcal{I}^\complement}) \Pr(\boldsymbol{Z}_1 < \boldsymbol{Z}_2)\end{aligned}$$

where $(\boldsymbol{Z}_1, \boldsymbol{Z}_2) \sim \mathrm{CSN}_{2d,d(n-2)}(\bar{\boldsymbol{\Sigma}}_{\mathcal{I}\mathcal{I}}, \boldsymbol{\varrho}_{\mathcal{I}^\complement,\mathcal{I}^\complement}, \bar{\boldsymbol{\Sigma}}_{\mathcal{I}^\complement\mathcal{I}^\complement;\mathcal{I}})$. Formula (4.25) follows by noting that

$$\boldsymbol{Z}_1 - \boldsymbol{Z}_2 = \begin{pmatrix}\boldsymbol{I}_d & -\boldsymbol{I}_d\end{pmatrix}\begin{pmatrix}\boldsymbol{Z}_1 \\ \boldsymbol{Z}_2\end{pmatrix}$$

and by applying (4.4)

To compute formula (4.26), we repeat the same procedure. With $\mathcal{D}(\boldsymbol{a}_1, \dots, \boldsymbol{a}_n) :=$

$\prod(-\infty, \boldsymbol{a}_i]$ we obtain

$$
\begin{aligned}
&\Pr(\boldsymbol{X}_j \leq \boldsymbol{x}_1,\ \boldsymbol{X}_n \leq \boldsymbol{x}_2, R_j^{\text{CR}} = 1, R_n^{\text{CR}} = 1) \\
&= \Pr(\boldsymbol{X}_j \leq \boldsymbol{x}_1,\ \boldsymbol{X}_n \leq \boldsymbol{x}_2, \boldsymbol{X}_j > \boldsymbol{M}_{j-1}, \boldsymbol{X}_n > \boldsymbol{M}_{j+1:n-1}\boldsymbol{X}_i, \boldsymbol{X}_j < \boldsymbol{X}_n) \\
&= \int_{(-\boldsymbol{\infty}, \boldsymbol{x}_2]} \sum_{\mathcal{J} \subseteq \{1,\dots,d\}} \left(\int_{\mathcal{D}(\boldsymbol{z}_{n\mathcal{J}}, \boldsymbol{x}_{1\bar{\mathcal{J}}})} \Pr(\boldsymbol{M}_{j-1} < \boldsymbol{z}_j, \boldsymbol{M}_{j+1:n-1} < \boldsymbol{z}_n | \boldsymbol{X}_i = \boldsymbol{z}_i, i \in \mathcal{I}) \right. \\
&\qquad \left. \cdot \phi_{2d}(\boldsymbol{z}_j, \boldsymbol{z}_n; \bar{\boldsymbol{\Sigma}}_{\mathcal{I}\mathcal{I}}) \mathrm{d}\boldsymbol{z}_j \right) \mathbb{1}(\boldsymbol{z}_{n\mathcal{J}} < \boldsymbol{x}_{1\mathcal{J}}, \boldsymbol{z}_{n\bar{\mathcal{J}}} > \boldsymbol{x}_{1\bar{\mathcal{J}}}) \mathrm{d}\boldsymbol{z}_n \\
&= \sum_{\mathcal{J} \subseteq \{1,\dots,d\}} \int_{(-\boldsymbol{\infty}, \boldsymbol{x}_{1\mathcal{J}}]} \int_{(\boldsymbol{x}_{1\mathcal{J}}, \boldsymbol{x}_{2\bar{\mathcal{J}}}]} \int_{\mathcal{D}(\boldsymbol{z}_{n\mathcal{J}}, \boldsymbol{x}_{1\bar{\mathcal{J}}})} \Phi_{d(n-2)}(\boldsymbol{\varrho}_{\mathcal{I}^\complement, \mathcal{I}^\complement} \boldsymbol{z}; \bar{\boldsymbol{\Sigma}}_{\mathcal{I}^\complement \mathcal{I}^\complement; \mathcal{I}}) \\
&\qquad \cdot \phi_{2d}(\boldsymbol{z}_j, \boldsymbol{z}_n; \bar{\boldsymbol{\Sigma}}_{\mathcal{I}\mathcal{I}}) \mathrm{d}\boldsymbol{z}_j \mathrm{d}\boldsymbol{z}_n \\
&= \sum_{\mathcal{J} \subseteq \{1,\dots,d\}} \left(\int_{\mathcal{D}(\boldsymbol{x}_{1\mathcal{J}}, \boldsymbol{x}_{2\bar{\mathcal{J}}}, \boldsymbol{z}_{n\mathcal{J}}, \boldsymbol{x}_{1\bar{\mathcal{J}}})} \Phi_{d(n-2)}(\boldsymbol{\varrho}_{\mathcal{I}^\complement, \mathcal{I}^\complement} \boldsymbol{z}; \bar{\boldsymbol{\Sigma}}_{\mathcal{I}^\complement \mathcal{I}^\complement; \mathcal{I}}) \phi_{2d}(\boldsymbol{z}_j, \boldsymbol{z}_n; \bar{\boldsymbol{\Sigma}}_{\mathcal{I}\mathcal{I}}) \mathrm{d}\boldsymbol{z}_j \mathrm{d}\boldsymbol{z}_n \right. \\
&\qquad \left. - \int_{\mathcal{D}(\boldsymbol{x}_{1\mathcal{J}}, \boldsymbol{x}_{2\bar{\mathcal{J}}}, \boldsymbol{z}_{n\mathcal{J}}, \boldsymbol{x}_{1\bar{\mathcal{J}}})} \Phi_{d(n-2)}(\boldsymbol{\varrho}_{\mathcal{I}^\complement, \mathcal{I}^\complement} \boldsymbol{z}; \bar{\boldsymbol{\Sigma}}_{\mathcal{I}^\complement \mathcal{I}^\complement; \mathcal{I}}) \phi_{2d}(\boldsymbol{z}_j, \boldsymbol{z}_n; \bar{\boldsymbol{\Sigma}}_{\mathcal{I}\mathcal{I}}) \mathrm{d}\boldsymbol{z}_j \mathrm{d}\boldsymbol{z}_n \right) \\
&= \sum_{\mathcal{J} \subseteq \{1,\dots,d\}} \Phi_{d(n-2)}(\boldsymbol{0}; \bar{\boldsymbol{\Sigma}}_{\mathcal{I}^\complement \mathcal{I}^\complement; \mathcal{I}} + \boldsymbol{\varrho}_{\mathcal{I}^\complement, \mathcal{I}^\complement} \bar{\boldsymbol{\Sigma}}_{\mathcal{I}\mathcal{I}} \boldsymbol{\varrho}_{\mathcal{I}^\complement, \mathcal{I}^\complement}^\top) \\
&\qquad (\Pr(\boldsymbol{Z}_{1\mathcal{J}} < \boldsymbol{Z}_{2\mathcal{J}}, \boldsymbol{Z}_{1\bar{\mathcal{J}}} \leq \boldsymbol{x}_{1\bar{\mathcal{J}}}, \boldsymbol{Z}_{2\mathcal{J}} \leq \boldsymbol{x}_{1\bar{\mathcal{J}}}, \boldsymbol{Z}_{2\bar{\mathcal{J}}} \leq \boldsymbol{x}_{2\bar{\mathcal{J}}}) \\
&\qquad - \Pr(\boldsymbol{Z}_{1\mathcal{J}} < \boldsymbol{Z}_{2\mathcal{J}}, \boldsymbol{Z}_{1\bar{\mathcal{J}}} \leq \boldsymbol{x}_{1\bar{\mathcal{J}}}, \boldsymbol{Z}_{2\mathcal{J}} \leq \boldsymbol{x}_{1\bar{\mathcal{J}}}, \boldsymbol{Z}_{2\bar{\mathcal{J}}} \leq \boldsymbol{x}_{1\bar{\mathcal{J}}}))
\end{aligned}
$$

The first probability in the right hand side can be computed by noting that

$$
\begin{pmatrix} \boldsymbol{Z}_{1\mathcal{J}} - \boldsymbol{Z}_{1\mathcal{J}} \\ \boldsymbol{Z}_{1\bar{\mathcal{J}}} \\ \boldsymbol{Z}_{2\mathcal{J}} \\ \boldsymbol{Z}_{2\bar{\mathcal{J}}} \end{pmatrix} = \begin{pmatrix} \boldsymbol{I}_{\mathcal{J}} & \boldsymbol{0}_{\bar{\mathcal{J}}} & -\boldsymbol{I}_{\mathcal{J}} & \boldsymbol{0}_{\bar{\mathcal{J}}} \\ \boldsymbol{0}_{\mathcal{J}} & \boldsymbol{I}_{\bar{\mathcal{J}}} & \boldsymbol{0}_{\mathcal{J}} & \boldsymbol{0}_{\bar{\mathcal{J}}} \\ \boldsymbol{0}_{\mathcal{J}} & \boldsymbol{0}_{\bar{\mathcal{J}}} & \boldsymbol{I}_{\mathcal{J}} & \boldsymbol{0}_{\bar{\mathcal{J}}} \\ \boldsymbol{0}_{\mathcal{J}} & \boldsymbol{0}_{\bar{\mathcal{J}}} & \boldsymbol{0}_{\mathcal{J}} & \boldsymbol{I}_{\bar{\mathcal{J}}} \end{pmatrix} \begin{pmatrix} \boldsymbol{Z}_{1\mathcal{J}} \\ \boldsymbol{Z}_{1\bar{\mathcal{J}}} \\ \boldsymbol{Z}_{2\mathcal{J}} \\ \boldsymbol{Z}_{2\bar{\mathcal{J}}} \end{pmatrix}
$$

where $(\boldsymbol{Z}_1, \boldsymbol{Z}_2) = (\boldsymbol{Z}_{1\mathcal{J}}, \boldsymbol{Z}_{1\bar{\mathcal{J}}}, \boldsymbol{Z}_{2\mathcal{J}}, \boldsymbol{Z}_{2\bar{\mathcal{J}}}) \sim \text{CSN}_{2d,d(n-2)}(\bar{\boldsymbol{\Sigma}}_{\mathcal{I}\mathcal{I}}, \boldsymbol{\varrho}_{\mathcal{I}^\complement, \mathcal{I}^\complement}, \bar{\boldsymbol{\Sigma}}_{\mathcal{I}^\complement \mathcal{I}^\complement; \mathcal{I}})$ and by applying (4.4). The second probability is computed in the same way. □

Chapter 5

Conclusions

5.1 Discussion

It is clear why records represent an extremely appealing topic: every subject has its own records, from the financial one to the environmental, passing through the medical or the sport. Records are therefore of interest in a countless number of disciplines, however the literature on record values is not developed as other topics in probability and statistics. As already explained by Arnold, Balakrishnan, and Nagaraja (1998), the issues to be faced in studying records arise from two main aspects. Firstly, if we do a little step beyond the hypothesis of i.i.d. observations, the problem becomes too complex. The second issue arises from a statistical point of view, since rarity of records leads to very small samples to deal with.

The purpose of this thesis is to provide a new point of view and new techniques to tackle the topic of record values and record times. Starting from the approach of not knowing the position of a record in the sequence of records, we firstly provide both asymptotic and non-asymptotic results concerning the distribution of a record and also the joint distribution of two and three records. In particular, we show the link existing between the distribution of records and the distribution of the maximum of a set of random variables, as well as the asymptotic distribution of a record and the GEV distribution. By following the idea of the asymptotic distribution of multivariate maxima in the classical Extreme Value

Theory, we compute the limiting distribution function of records over a sequence of i.i.d. random variables with an *unkown* distribution function, provided it is continuous.

Topics discussed in Chapter 3 deal with records in higher dimensions. As we already pointed out, it is not clear in general how to define a multivariate record, since a natural ordering in $\mathbb{R}^d, d > 1$ does not exist. We considered the definition of *complete record* and we studied the case of an i.i.d. sequence of random vectors with independent components. Since it has been proved by Goldie and Resnick (1989) that the number of complete records in this framework is finite, we focused our attention on the distribution of the last complete record, which resulted in a very complicated expression. However, inspired by the fact that the probability of observing a new record after the first one, i.e. $\boldsymbol{X}_1$, tends very fastly to zero when the dimension increases, we prove that the distribution of the last complete record is approximated by the distribution of $\boldsymbol{X}_1$.

The second part of Chapter 3 is dedicated to study the properties of the arrival times process of complete records over sequences of random vectors with independent components. In particular, we show the Markovian property of such a process, by fournishing the transition probabilities. Clearly, this result extends what has been proved by Galambos (1987).

Our contributions in Chapter 4 have a remarkeable potential in practical applications, since we extend previous results to the case of a second order stationary gaussian process. First of all we deal with stationary sequences of random variables and we fournish exact distributions for single and joint records over time, as well as some results on the arrival times of records. Moreover, from an asymptotic point of view, we discuss the role of the *extremal index* for stationary sequences of gaussian processes in Extreme Value Theory, in the framework of record values. Notice that the extremal index can be estimated by using the entire sample, not just the record values, so that the problem of small samples in estimating record distributions is solved at least in this case. Finally, our analysis provides new results for *complete records* in the case of multivariate second order stationary gaussian random vectors with dependent components, by reaching the goal of computing exact distributions for both single and joint-in-time complete records. In the remarks we show how some of the results in the i.i.d. framework can be obtained as a particular case of our new findings.

5.2 Future work

The results presented in this thesis set the base for answering to new or still open questions. First, let us consider an i.i.d. sequence of random vectors with dependent components. Then, this dependence can be described by means of the so-called D-norm (see Dombry, Falk, and Zott (2015) and Zott (2016)). An asymptotic joint distribution in time for two complete records is still missing, even if the use of the D-norm should ease the computations. Furthermore, the expression computed for the joint distribution in time of two simple records has a really difficult form.

The problem of hardly manageable results has to be faced for complete records over multivariate gaussian processes as well. In this case it would be desirable to find more tractable expressions for the closed skew-normal distributions, probably imposing restrictions over the structure of the underlying gaussian process (Markovian structure, autoregressive, ARMA etc.). Moreover, our results deal only with one record and two records and computations become really hard when we increase this number.

For what concerns record times, it would be useful to show the properties of the process for an underlying sequence of dependent random variables. It has been proven by Galambos (1987), that the arrival times process of records is a Markov chain in the case of records over a sequence of i.i.d. random variables. The question is if it is possible to weaken these hypotheses to obtain the markovian property of the arrival time process of records.

Bibliography

Anderson, T. W. (1984). *An Introduction to Multivariate Statistical Analysis* (Second ed.). Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York.

Arnold, B. C., N. Balakrishnan, and H. N. Nagaraja (1998). *Records*. Wiley Series in Probability and Statistics. New York: Wiley.

Azzalini, A. (2013). *The Skew-Normal and Related Families*, Volume 3. Cambridge University Press.

Azzalini, A. and A. Bacchieri (2010). A prospective combination of phase ii and phase iii in drug development. *Metron 68*(3), 347–369.

Azzalini, A. and A. Capitanio (1999). Statistical applications of the multivariate skew normal distribution. *J. R. Statist. Soc. B 61*(3), 579–602.

Azzalini, A. and A. D. Valle (1996). The multivariate skew-normal distribution. *Biometrika 83*(4), 715–726.

Ballerini, R. and S. I. Resnick (1987). Records in the presence of a linear trend. *Adv. in Appl. Probab. 19*(4), 801–828.

Banerjee, S., B. P. Carlin, and A. E. Gelfand (2014). *Hierarchical Modeling and Analysis for Spatial Data*. Crc Press.

Barakat, H. and M. A. Elgawad (2017). Asymptotic behavior of the joint record values, with applications. *Statistics & Probability Letters 124*, 13–21.

Brockwell, P. J. and R. A. Davis (2013). *Time Series: Theory and Methods*. Springer Science & Business Media.

Chandler, K. N. (1952). The distribution and frequency of record values. *J. R. Statist. Soc. B 14*(2), 220–228.

Chernick, M. R. et al. (1981). A limit theorem for the maximum of autoregressive processes with uniform marginal distributions. *The Annals of Probability 9*(1), 145–149.

Coles, S., J. Bawa, L. Trenner, and P. Dorazio (2001). *An Introduction to Statistical Modeling of Extreme Values*, Volume 208. Springer.

Cressie, N. and C. K. Wikle (2015). *Statistics for Spatio-Temporal Data*. John Wiley & Sons.

Dombry, C., M. Falk, and M. Zott (2015). On functional records and champions. *Journal of Theoretical Probability*, 1–26.

Falk, M., A. K. Chokami, and S. Padoan (2018). Some results on joint record events. *Statistics & Probability Letters 135*, 11–19.

Falk, M., J. Hüsler, and R.-D. Reiss (2011). *Laws of Small Numbers: Extremes and Rare Events* (3 ed.). Basel: Springer.

Falk, M., A. Khorrami Chokami, and S. A. Padoan (2018). On multivariate records from random vectors with independent components. *J. Appl. Probab. 55*(1), 43–53.

Galambos, J. (1987). *The Asymptotic Theory of Extreme Order Statistics* (2 ed.). Malabar: Krieger.

Genton, M. G. (2004). *Skew-Elliptical Distributions and Their Applications: A Journey Beyond Normality*. CRC Press.

Goldie, C. M. and S. I. Resnick (1989). Records in a partially ordered set. *Ann. Probab. 17*(2), 678–699.

Goldie, C. M. and S. I. Resnick (1995). Many multivariate records. *Stochastic Process. Appl. 59*(2), 185–216.

Haiman, G. (1987). Almost sure asymptotic behavior of the record and record time sequences of a stationary gaussian process. In *Mathematical Statistics and Probability Theory*, pp. 105–120. Springer.

Haiman, G., N. Mayeur, V. Nevzorov, and M. L. Puri (1998). Records and 2-block records of 1-dependent stationary sequences under local dependence. In *Annales de l'Institut Henri Poincare (B) Probability and Statistics*, Volume 34, pp. 481–503. Elsevier.

Hsing, T., J. Hüsler, R.-D. Reiss, et al. (1996). The extremes of a triangular array of normal random variables. *The Annals of Applied Probability 6*(2), 671–686.

Kurowicka, D. and R. M. Cooke (2006). *Uncertainty Analysis With High Dimensional Dependence Modelling*. John Wiley & Sons.

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

Lindgren, G. (2012). *Stationary Stochastic Processes: Theory and Applications*. CRC Press.

O'Brien, G. L. (1987). Extreme values for stationary and markov sequences. *The Annals of Probability*, 281–291.

Rényi, A. (1962). Théorie des éléments saillants d'une suite d'observations. *Ann. scient. Univ. de Clermont. Mathématiques 8*(2), 7–13.

Resnick, S. I. (1973). Limit laws for record values. *Stochastic Processes and Their Applications 1*(1), 67–82.

Resnick, S. I. (1987). *Extreme Values, Regular Variation, and Point Processes*, Volume 4 of *Applied Probability*. New York: Springer. First Printing.

Rootzén, H. (1988). Maxima and exceedances of stationary markov chains. *Advances in applied probability 20*(2), 371–390.

Zott, M. (2016). *Extreme Value Theory in Higher Dimensions. Max-Stable Processes and Multivariate Records*. Ph. D. thesis, University of Würzburg.