On multivarate records over sequences of random vectors with Marshall-Olkin dependence of components

Record multivariati su successioni di vettori aleatori con dipendenza di tipo Marshall-Olkin

M. Falk, A. Khorrami and S. A. Padoan

Abstract Let $X_1, X_2, \dots$ be independent and identically distributed random variables on the real line with a joint continuous distribution function F. The stochastic behavior of the sequence of subsequent records is well known. Falk, Khorrami and Padoan (2018) considered multivariate records over sequences of random vectors with independent components, providing results on the arrival times process of the records. In this work, we study the case of records over sequences of random vectors where the dependence among their components is of the Marshall-Olkin type.

Abstract *Sia* $X_1, X_2, \dots$ *una successione di variabili aleatorie indipendenti ed identicamente distribuite sulla retta reale, con distribuzione continua* F*. La succesione dei record in questo caso stata ampiamente studiata in passato. Falk, Khorrami and Padoan (2018) hanno considerato il caso di record multivariati su successioni di vettori aleatori indipendenti, con componenti indipendenti, fornendo risultati riguardanti il processo dei tempi d'arrivo dei record. In questo lavoro, invece, gli autori studiano i record multivariati su successioni di vettori aleatori indipendenti, ma con dipendenza fra le loro componenti di tipo Marshall-Olkin.*

M. Falk
Institute of Mathematics, University of Würzburg, Am Hubland, D-97074, Würzburg, Germany, e-mail: michael.falk@uni-wuerzburg.de

A. Khorrami and S. A. Padoan
Department of Decision Sciences, Bocconi University of Milan, via Roentgen 1, 20136 Milano, Italy e-mail: amir.khorrami@phd.unibocconi.it, e-mail: simone.padoan@unibocconi.it

1 Introduction

Let us consider a sequence of independent and identically distributed (iid) random variables $X_1, X_2, \dots$. The rv X_m is called a *record* if $X_m > \max(X_1, \dots, X_{m-1})$, $m \geq 2$. Clearly, X_1 is a record. Records among a sequence of iid rv on the real line have been investigated extensively over the past decades, see [4, Sections 6.2 and 6.3] and [1]. It is, for example, well known that in the univariate case the indicator functions $I_m^{\mathrm{R}} = \mathbb{1}\left(X_m > \max(X_1, \dots, X_{m-1})\right)$, $m \in \mathbb{N}$, are independent, see, e.g., [4, Lemma 6.3.3]. This implies in particular for integers $i \neq k$

$$\Pr(I_j^{\mathrm{R}} = 1, I_k^{\mathrm{R}} = 1) = \Pr(I_j^{\mathrm{R}} = 1)\Pr(I_k^{\mathrm{R}} = 1) = j^{-1}k^{-1}. \tag{1}$$

More recently, [2] found new results concerning records over sequences of random variables, without knowing their position in the sequence of records.

Let $\boldsymbol{X}_1, \boldsymbol{X}_2, \dots$ be iid random vectors (rv) in $\mathbb{R}^d$. Put for integers $j < k$

$$\boldsymbol{M}_j := \max_{1 \leq i \leq j} \boldsymbol{X}_i, \quad \boldsymbol{M}_j^k := \max_{j \leq i \leq k} \boldsymbol{X}_i,$$

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

Trying to generalize the concept of records to rv $\boldsymbol{X}_1, \boldsymbol{X}_2, \dots$ in $\mathbb{R}^d$, the question which arises is how to define records in higher dimensions. Various definitions of records have been proposed in the past, see for example [5]. In this work, we consider the case of complete records: a rv $\boldsymbol{X}_m$ is a *complete* record if each component is a record itself, i.e., $\boldsymbol{X}_m > \boldsymbol{M}_{m-1}$.

Multivariate records have not been discussed that extensively, yet they have been approached by e. g. [6] or [1, Chapter 8]. For supplementary material on multivariate and functional records we refer to the thesis by [7].

We denote the event that $\boldsymbol{X}_m$ is a complete record by the indicator function $I_m^{\mathrm{CR}} := \mathbb{1}(\boldsymbol{X}_m > \boldsymbol{M}_{m-1})$. Results concerning complete records over sequences of random vectors with independent components can be found in [3].

This paper investigates, amon others, the problem, which results for univariate records *do not* carry over to the multivariate case.

2 Marshall-Olkin

We have independence of record events in the univariate case and in the multivariate case with complete margins. One might conjecture that this is true, if we consider a convex combination of these two. The Marshall-Olkin standard max-stable distributions exactly provide this approach. Clearly, when the elements of $\boldsymbol{X}$ are complete dependent, then the probability that two joint records take place is equal to (1). We saw that this is also the case when the elements of $\boldsymbol{X}$ are independent. Therefore, one can conjecture that it also holds true when the elements of $\boldsymbol{X}$ are dependent. We show with a counter-example that this is not true.

For simplicity assume $d = 2$. Suppose that $\boldsymbol{\eta}$ is a two-dimensional Marshall-Olkin distribution function:

$$\Pr(\boldsymbol{\eta} \leq \boldsymbol{x}) = \exp\left(-\{\lambda \left\|\boldsymbol{x}\right\|_\infty + (1-\lambda)\left\|\boldsymbol{x}\right\|_1\}\right), \qquad \boldsymbol{x} \leq \boldsymbol{0} \in \mathbb{R}^2, \tag{2}$$

where $\lambda \in [0,1]$ is a dependence parameter. The comoponents of $\boldsymbol{\eta}$ are completely dependent when $\lambda = 1$, while they are independent when $\lambda = 0$.

Theorem 1. *Let $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \dots$ be iid rv with distribution function* (2)*. Then,*

$$\Pr(I_n^{CR} = 1) = \frac{1}{n(2-\lambda)}\left(\frac{2(1-\lambda)}{n} + \lambda\right) \tag{3}$$

and

$$\Pr(I_{j,k}^{CR} = 1) = \frac{1}{j^2k(2-\lambda)}\Bigg(\frac{2(1-\lambda)}{k(2-\lambda)}\frac{(2-\lambda)k+(2-\lambda)(1-\lambda)j-\lambda(1-j)k}{k+(1-\lambda)j} + \lambda(2(1-\lambda)+j\lambda)\Bigg). \tag{4}$$

Remark 1. When $\lambda = 0$ equation (4) becomes

$$\Pr(I_{j,k}^{\mathrm{CR}} = 1) = \frac{1}{j^2k^2},$$

while

$$\Pr(I_{j,k}^{\mathrm{CR}} = 1) = \frac{1}{jk},$$

when $\lambda = 1$, as expected.

Now, note that

$$\Pr(I_j^{\mathrm{CR}} = 1)\Pr(I_k^{\mathrm{CR}} = 1) = \frac{1}{jk(2-\lambda)^2}\left(\frac{2(1-\lambda)}{j} + \lambda\right)\left(\frac{2(1-\lambda)}{k} + \lambda\right)$$

and from this and (4) we have

$$\begin{aligned}&\Pr(I_{j,k}^{\mathrm{CR}} = 1) - \Pr(I_j^{\mathrm{CR}} = 1)\Pr(I_k^{\mathrm{CR}} = 1)\\ &= \frac{(\lambda-1)\lambda(k^2(\lambda(j-2)+2) - (\lambda-1)jk(\lambda(j-2)+2) + 2(\lambda-1)(j-1)j)}{(\lambda-2)^2 j^2k^2((\lambda-1)j-k)}\end{aligned}$$

and therefore the events $(I_j^{\mathrm{CR}} = 1)$ and $(I_k^{\mathrm{CR}} = 1)$ are not independent.

Proof (Theorem 1).

$$\begin{aligned}
\Pr(I_n^{\mathrm{CR}}=1) = \Pr(\boldsymbol{\eta}_n > \boldsymbol{M}_{n-1}) &= \int_{(-\infty,0]^2} \Pr(\boldsymbol{\eta} < (n-1)\boldsymbol{x}) \mathrm{d}\Pr_{\boldsymbol{\eta}}(\boldsymbol{x}) \\
&= \int_{-\infty}^{0}\int_{-\infty}^{x_2} (1-\lambda)\,\mathrm{e}^{n(x_1+(1-\lambda)x_2)}\,\mathrm{d}x_1\mathrm{d}x_2 \\
&+ \int_{-\infty}^{0}\int_{x_2}^{0} (1-\lambda)\,\mathrm{e}^{n(x_2+(1-\lambda)x_1)}\,\mathrm{d}x_1\mathrm{d}x_2 \\
&+ \int_{-\infty}^{0} \lambda\,\mathrm{e}^{n(2-\lambda)x}\,\mathrm{d}x \\
&= \frac{1-\lambda}{2-\lambda}\frac{1}{n^2} + \frac{1}{n^2} - \frac{1}{2-\lambda}\frac{1}{n^2} + \frac{\lambda}{2-\lambda}\frac{1}{n}.
\end{aligned}$$

For what concerns the probability of the joint events we have

$$\begin{aligned}
\Pr(I_j^{\mathrm{CR}}=1, I_k^{\mathrm{CR}}=1) &= \Pr(\boldsymbol{\eta}_j > \boldsymbol{M}_{j-1}, \boldsymbol{\eta}_k > \boldsymbol{M}_{k-1}) \\
&= \Pr\left(\boldsymbol{\eta}_j > \frac{\boldsymbol{\eta}_1}{j-1}, \boldsymbol{\eta}_k > \max\left(\boldsymbol{\eta}_j, \frac{\boldsymbol{\eta}_2}{k-j-1}\right)\right) \\
&= \int_{(-\infty,0]^d}\int_{(-\infty,\boldsymbol{y}]} \Pr\left(\boldsymbol{\eta}_1 < (j-1)\boldsymbol{x}, \boldsymbol{\eta}_2 < (k-j-1)\boldsymbol{y}\right)\mathrm{d}\Pr_{\boldsymbol{\eta}}(\boldsymbol{x})\mathrm{d}\Pr_{\boldsymbol{\eta}}(\boldsymbol{y}) \\
&= \int_{(-\infty,0]^2}\int_{(-\infty,\boldsymbol{y}]} \mathrm{e}^{(j-1)\|\boldsymbol{x}\|_\lambda + (k-j-1)\|\boldsymbol{y}\|_\lambda}\,\mathrm{d}\Pr_{\boldsymbol{\eta}}(\boldsymbol{x})\mathrm{d}\Pr_{\boldsymbol{\eta}}(\boldsymbol{y}) \\
&= \int_{-\infty}^{0}\int_{-\infty}^{y_1} \mathrm{e}^{(k-j-1)\|\boldsymbol{y}\|_\lambda} I(\boldsymbol{y})\mathrm{d}\Pr_{\boldsymbol{\eta}}(\boldsymbol{y}) + \int_{-\infty}^{0}\int_{-\infty}^{y_2} \mathrm{e}^{(k-j-1)\|\boldsymbol{y}\|_\lambda} I(\boldsymbol{y})\mathrm{d}\Pr_{\boldsymbol{\eta}}(\boldsymbol{y}) \\
&+ \int_{-\infty}^{0} \mathrm{e}^{(k-j-1)\|\boldsymbol{y}\|_\lambda} I(\boldsymbol{y})\mathrm{d}\Pr_{\boldsymbol{\eta}(\boldsymbol{y})} = I_1 + I_2 + I_3,
\end{aligned}$$

where

$$I(\boldsymbol{y}) = \int_{(-\infty,\boldsymbol{y}]} \mathrm{e}^{(j-1)\|\boldsymbol{x}\|_\lambda}\,\mathrm{d}\Pr_{\boldsymbol{\eta}}(\boldsymbol{x}).$$

Computation of I_1. In the case $y_1 < y_2$, we have

$$\begin{aligned}
I(\boldsymbol{y}) &= \int_{-\infty}^{y_1}\int_{x_1}^{y_2} (1-\lambda)\,\mathrm{e}^{j(x_1+(1-\lambda)x_2)}\,\mathrm{d}x_2\mathrm{d}x_1 + \int_{-\infty}^{y_1}\int_{-\infty}^{x_1} (1-\lambda)\,\mathrm{e}^{j(x_2+(1-\lambda)x_1)}\,\mathrm{d}x_2\mathrm{d}x_1 \\
&+ \int_{-\infty}^{y_1} \lambda\,\mathrm{e}^{j(2-\lambda)}\,\mathrm{d}x = I_{1,1} + I_{1,2} + I_{1,3}.
\end{aligned}$$

Integrals $I_{1,1}$, $I_{1,2}$ and $I_{1,3}$ are equal to

$$\begin{aligned}
I_{1,1} &= (1-\lambda)\int_{-\infty}^{y_1} \mathrm{e}^{jx_1}\,\frac{e^{j(1-\lambda)y_2} - \mathrm{e}^{j(1-\lambda)x_1}}{j(1-\lambda)}\mathrm{d}x_1 = \frac{1}{j}\left(\mathrm{e}^{j(1-\lambda)y_2}\,\frac{\mathrm{e}^{jy_1}}{j} - \frac{\mathrm{e}^{j(2-\lambda)y_1}}{j(2-\lambda)}\right), \\
I_{1,2} &= (1-\lambda)\int_{-\infty}^{y_1} \mathrm{e}^{j(1-\lambda)x_1}\,\frac{\mathrm{e}^{j\lambda x_1}}{j}\mathrm{d}x_1 = \frac{1-\lambda}{j^2}\,\frac{\mathrm{e}^{j(2-\lambda)y_1}}{j(2-\lambda)} \\
I_{1,3} &= \lambda\,\frac{\mathrm{e}^{j(2-\lambda)y_1}}{j(2-\lambda)}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
I_1 &= \int_{-\infty}^{0}\int_{-\infty}^{y_1} \frac{1}{j^2}\,\mathrm{e}^{j(y_1+(1-\lambda)y_2)} - \frac{1-j}{j^2(2-\lambda)}\lambda\,\mathrm{e}^{j(2-\lambda)y_1}\,\mathrm{e}^{(k-j)(y_1+(1-\lambda)y_2}\,\mathrm{d}y_1\mathrm{d}y_2 \\
&= \frac{1-\lambda}{j^2}\int_{-\infty}^{0}\mathrm{e}^{k(1-\lambda)y_2}\,\frac{\mathrm{e}^{ky_2}}{k} - \frac{\lambda}{2-\lambda}\,\frac{1-j}{k+(1-\lambda)j}\,\mathrm{e}^{(2-\lambda)k}\,\mathrm{d}y_2 \\
&= \frac{1-\lambda}{(2-\lambda)^2}\,\frac{1}{j^2k^2}\,\frac{(2-\lambda)k+(2-\lambda)(1-\lambda)j-\lambda(1-j)k}{k+(1-\lambda)j}.
\end{aligned}$$

The computation for I_2 is similar to that of I_1 and we obtain $I_2 = I_1$.

Computation of I_3. In the case $y_1 = y_2$, we have

$$\begin{aligned}
I(\boldsymbol{y}) &= 2\int_{-\infty}^{y}\int_{-\infty}^{x_2}(1-\lambda)\,\mathrm{e}^{j(x_1+(1-\lambda)x_2)}\,\mathrm{d}x_1\mathrm{d}x_2 + \int_{-\infty}^{y}\lambda\,\mathrm{e}^{j(2-\lambda)}\,\mathrm{d}x \\
&= 2\frac{1-\lambda}{j^2(2-\lambda)}\,\mathrm{e}^{j(2-\lambda)y} + \lambda\,\frac{\mathrm{e}^{j(2-\lambda)y}}{j(2-\lambda)} = 2\frac{1-(1-j)\lambda}{j^2(2-\lambda)}\,\mathrm{e}^{j(2-\lambda)y}.
\end{aligned}$$

Therefore,

$$I_3 = \int_{-\infty}^{0}\lambda\left(2\frac{1-\lambda}{j^2(2-\lambda)} + \frac{\lambda}{j(2-\lambda)}\right)\mathrm{e}^{k(2-\lambda)}\,\mathrm{d}x = \lambda\,\frac{2(1-\lambda)+\lambda j}{(2-\lambda)}\,\frac{1}{j^2k}.$$

Finally

$$\Pr(I_j^{\mathrm{CR}} = 1, I_k^{\mathrm{CR}} = 1) = 2I_1 + I_3.$$

It is still an open question how to show this result for a sequence $\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_n$, without assuming a specific dependence form.

Note that from the previous formulas we can deduce $\Pr(I_n^{\mathrm{CR}} = 1\,i.o.)$:

- Case $\lambda = 1$. Since the I_n^{R}'s are independent and $\sum_{n=2}^{\infty}\frac{1}{n} = \infty$, then
$$\Pr(I_n^{\mathrm{CR}} = 1\,i.o.) = 1$$
by the second Borel-Cantelli Lemma;
- Case $\lambda = 0$. Since the I_n^{CR}'s are independent and $\sum_{n=2}^{\infty}\frac{1}{n^d} < \infty$ if $d > 1$, then
$$\Pr(I_n^{\mathrm{CR}} = 1\,i.o.) = 0$$
by the first Borel-Cantelli Lemma,
- Case $\lambda \in (0,1)$. Call $N = \sum_{n=2}^{\infty} I_n^{\mathrm{CR}}$, then
$$\mathrm{e}[N] = \sum_{n=2}^{\infty}\mathrm{e}[I_n^{\mathrm{CR}}] = \sum_{n=2}^{\infty}\Pr(I_n^{\mathrm{CR}} = 1) = \sum_{n=0}^{\infty} 2\frac{1-\lambda}{2-\lambda}\frac{1}{n^2} + \frac{\lambda}{2-\lambda}\frac{1}{n} = \infty.$$

These results are consistent with Theorem 5.3 in [5].

References

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

[2] Falk, M., A. Khorrami, and S. A. Padoan (2017). Some results on joint record events. *Statistics and Probability Letters 135*(11–19).

[3] Falk, M., A. Khorrami, and S. A. Padoan (2018). On multivariate records from random vectors with independent components. *Journal of Applied Probability 135*(11–19).

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

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

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

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