On Limit Laws for Multivariate Record-breaking Processes

Ao Sun

Joint work with Jim Fill

Department of Applied Mathematics and Statistics
Johns Hopkins University

AofA2023@Taipei
June 30, 2023

Overview

1. Background
   - Classical univariate records
   - Multivariate records
2. Limit laws for multivariate record-breaking
   - Motivation
   - Record-breaking for general d
3. Record breaking in $\mathbb{R}^2$
   - Unit square
   - Unit isosceles right triangle
4. Perfect sampling algorithms

Table of Contents

Univariate records

Definition

Let $(X_n)_{n\geq 1}$ be a sequence of i.i.d. real-valued random variables with a common cumulative distribution function F. An observation X_n is called a *record* if its value exceeds that of all previous observations, i.e., $X_n > X_m$ for all $1 \leq m \leq n-1$.

To simplify our models by avoiding ties, we assume that F is a continuous function hereafter.

Univariate records

Definition

Let $(X_n)_{n\geq 1}$ be a sequence of i.i.d. real-valued random variables with a common cumulative distribution function F. An observation X_n is called a *record* if its value exceeds that of all previous observations, i.e., $X_n > X_m$ for all $1 \leq m \leq n-1$.

To simplify our models by avoiding ties, we assume that F is a continuous function hereafter.

Based on the fundamental work about univariate records due to Alfréd Rényi, their probabilistic properties are extensively studied in the literature; a rather encyclopedic treatment of this subject can be found in Arnold, Balakrishnan, and Nagaraja (1998).

Motivation

Dominance relation

For two vectors $\mathbf{x} = (x_1, \ldots, x_d)$ and $\mathbf{y} = (y_1, \ldots, y_d)$ in $\mathbb{R}^d$, we write $\mathbf{x} \prec \mathbf{y}$ to indicate that $x_i < y_i$ for all $i \in \{1, \ldots, d\}$. The notation $\mathbf{x} \succ \mathbf{y}$ means that $\mathbf{y} \prec \mathbf{x}$.

Motivation

Dominance relation

For two vectors $\mathbf{x} = (x_1, \ldots, x_d)$ and $\mathbf{y} = (y_1, \ldots, y_d)$ in $\mathbb{R}^d$, we write $\mathbf{x} \prec \mathbf{y}$ to indicate that $x_i < y_i$ for all $i \in \{1, \ldots, d\}$. The notation $\mathbf{x} \succ \mathbf{y}$ means that $\mathbf{y} \prec \mathbf{x}$.

Consider a sequence of i.i.d. random vectors $(\mathbf{X}^{(n)})_{n \geq 1}$ taking values in $\mathbb{R}^d$. If we go beyond one-dimensional Euclidean space, a natural way to define multivariate records would be that a multivariate record (or simply record) occurs at epoch n if $\mathbf{X}^{(n)} \succ \mathbf{X}^{(m)}$ for all $1 \leq m \leq n - 1$.

Motivation

Suppose the components of each observation are also independent, each having a continuous distribution function F_i, $i = 1, \ldots, d$. It is elementary to see from the previous definition that

$$\mathbb{P}(\mathbf{X}^{(n)} \text{ is a multivariate record}) = \left(\frac{1}{n}\right)^d = \frac{1}{n^d}.$$

Motivation

Suppose the components of each observation are also independent, each having a continuous distribution function F_i, $i = 1, \ldots, d$. It is elementary to see from the previous definition that

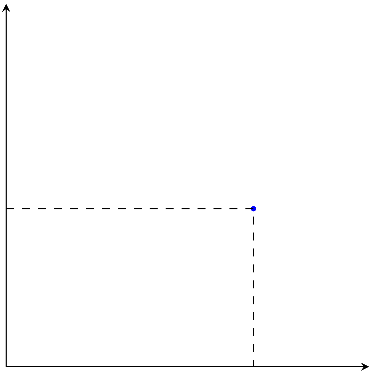
$$\mathbb{P}(\mathbf{X}^{(n)} \text{ is a multivariate record}) = \left(\frac{1}{n}\right)^d = \frac{1}{n^d}.$$

If $d \geq 2$, the first Borel-Cantelli lemma shows that there are at most *finitely many* such records with probability one. Hence the asymptotic results concerning such records are not that interesting.

An illustration of Pareto records in $\mathbb{R}^2_{\geq 0}$

In order to have a model that produces sufficiently many records, we consider *Pareto records* (formal definition later) in our project.

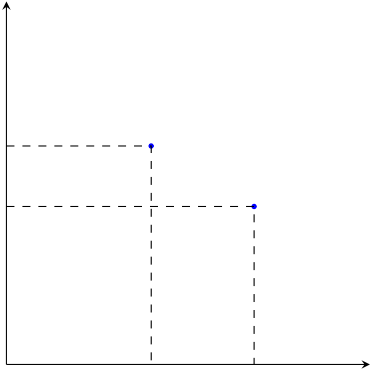
Bivariate Pareto records in $\mathbb{R}^2_{\geq 0}$



An illustration of Pareto records in $\mathbb{R}^2_{\geq 0}$

In order to have a model that produces sufficiently many records, we consider *Pareto records* (formal definition later) in our project.

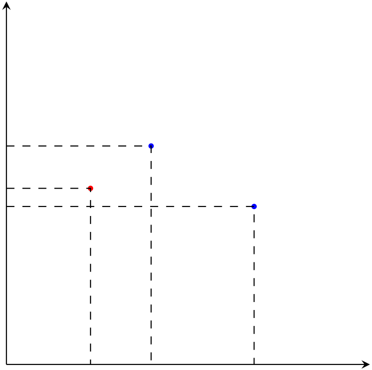
Bivariate Pareto records in $\mathbb{R}^2_{\geq 0}$



An illustration of Pareto records in $\mathbb{R}^2_{\geq 0}$

In order to have a model that produces sufficiently many records, we consider *Pareto records* (formal definition later) in our project.

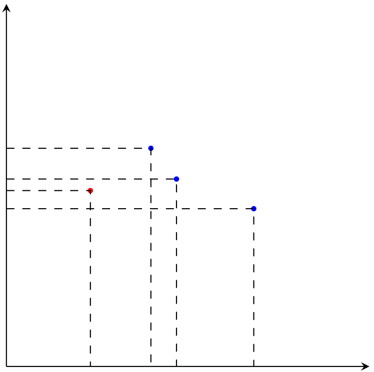
Bivariate Pareto records in $\mathbb{R}^2_{\geq 0}$



An illustration of Pareto records in $\mathbb{R}^2_{\geq 0}$

In order to have a model that produces sufficiently many records, we consider *Pareto records* (formal definition later) in our project.

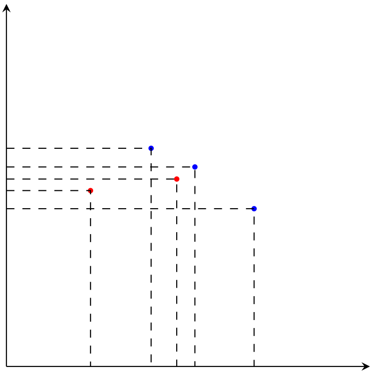
Bivariate Pareto records in $\mathbb{R}^2_{\geq 0}$



An illustration of Pareto records in $\mathbb{R}^2_{\geq 0}$

In order to have a model that produces sufficiently many records, we consider *Pareto records* (formal definition later) in our project.

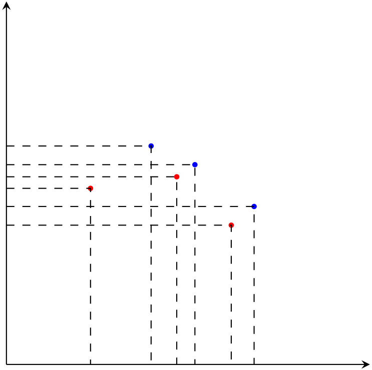
Bivariate Pareto records in $\mathbb{R}^2_{\geq 0}$



An illustration of Pareto records in $\mathbb{R}^2_{\geq 0}$

In order to have a model that produces sufficiently many records, we consider *Pareto records* (formal definition later) in our project.

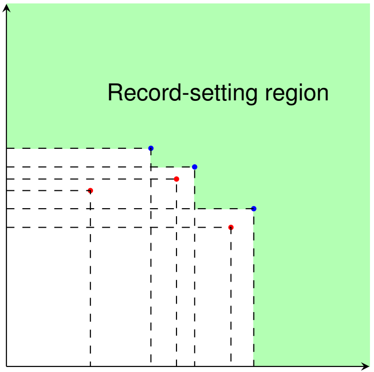
Bivariate Pareto records in $\mathbb{R}^2_{\geq 0}$



An illustration of Pareto records in $\mathbb{R}^2_{\geq 0}$

In order to have a model that produces sufficiently many records, we consider *Pareto records* (formal definition later) in our project.

Bivariate Pareto records in $\mathbb{R}^2_{\geq 0}$



Multivariate Pareto records: definitions

Definition

Let $(\mathbf{X}^{(n)})_{n\geq 1}$ be a sequence of i.i.d. random vectors in $\mathbb{R}^d$.

Multivariate Pareto records: definitions

Definition

Let $(\mathbf{X}^{(n)})_{n\geq 1}$ be a sequence of i.i.d. random vectors in $\mathbb{R}^d$.

1. The nth observation $\mathbf{X}^{(n)}$ is called a *Pareto record* (or *record*) if $\mathbf{X}^{(n)} \not\prec \mathbf{X}^{(i)}$ for $1 \leq i \leq n-1$.

Multivariate Pareto records: definitions

Definition

Let $(\mathbf{X}^{(n)})_{n\geq 1}$ be a sequence of i.i.d. random vectors in $\mathbb{R}^d$.

1. The nth observation $\mathbf{X}^{(n)}$ is called a *Pareto record* (or *record*) if $\mathbf{X}^{(n)} \not\prec \mathbf{X}^{(i)}$ for $1 \leq i \leq n-1$.
2. For $1 \leq i \leq n$, the ith observation $\mathbf{X}^{(i)}$ is called a *remaining record* (or *current record*, or *maximum*) at epoch n if $\mathbf{X}^{(i)} \not\prec \mathbf{X}^{(j)}$ for all $1 \leq j \leq n$.

Multivariate Pareto records: definitions

Definition

Let $(\mathbf{X}^{(n)})_{n\geq 1}$ be a sequence of i.i.d. random vectors in $\mathbb{R}^d$.

1. The nth observation $\mathbf{X}^{(n)}$ is called a *Pareto record* (or *record*) if $\mathbf{X}^{(n)} \not\prec \mathbf{X}^{(i)}$ for $1 \leq i \leq n-1$.
2. For $1 \leq i \leq n$, the ith observation $\mathbf{X}^{(i)}$ is called a *remaining record* (or *current record*, or *maximum*) at epoch n if $\mathbf{X}^{(i)} \not\prec \mathbf{X}^{(j)}$ for all $1 \leq j \leq n$.
3. For $1 \leq k \leq n$, the nth observation $\mathbf{X}^{(n)}$ is said to *break* k Pareto records if there exist precisely k observations $\mathbf{X}^{(j)}$ with $1 \leq j \leq n-1$ such that $\mathbf{X}^{(j)}$ is a current record at epoch $n-1$ and $\mathbf{X}^{(j)} \prec \mathbf{X}^{(n)}$.

Multivariate Pareto records: notations

Key notations

1. Let $r_n(d)$ denote the number of remaining Pareto records in $\mathbb{R}^d$ at epoch n; together these form the stochastic process $r(d) = (r_n(d))_{n \geq 1}$.

Multivariate Pareto records: notations

Key notations

1. Let $r_n(d)$ denote the number of remaining Pareto records in $\mathbb{R}^d$ at epoch n; together these form the stochastic process $r(d) = (r_n(d))_{n\geq 1}$.
2. Let $K_n(d)$ denote the number of remaining records in $\mathbb{R}^d$ at epoch $n-1$ broken by the nth observation $\mathbf{X}^{(n)}$. If $\mathbf{X}^{(n)}$ is not a record, then set $K_n(d) = -1$. We call the process $K(d) := (K_n(d))_{n\geq 1}$ the *record-breaking process*.

Multivariate Pareto records: notations

Key notations

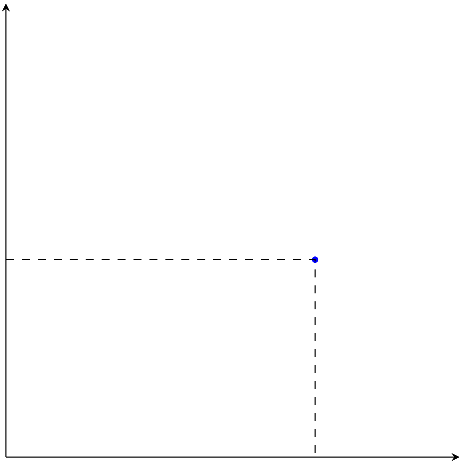
1. Let $r_n(d)$ denote the number of remaining Pareto records in $\mathbb{R}^d$ at epoch n; together these form the stochastic process $r(d) = (r_n(d))_{n\geq 1}$.
2. Let $K_n(d)$ denote the number of remaining records in $\mathbb{R}^d$ at epoch $n-1$ broken by the nth observation $\mathbf{X}^{(n)}$. If $\mathbf{X}^{(n)}$ is not a record, then set $K_n(d) = -1$. We call the process $K(d) := (K_n(d))_{n\geq 1}$ the *record-breaking process*.

When there is no confusion about the dimension d of the ambient space for $r(d)$ and $K(d)$, we will suppress the dependence on d and write these processes as $r = (r_n)_{n\geq 1}$ and $K = (K_n)_{n\geq 1}$ instead.

A graphical illustration for notations

$r_1 = 1$

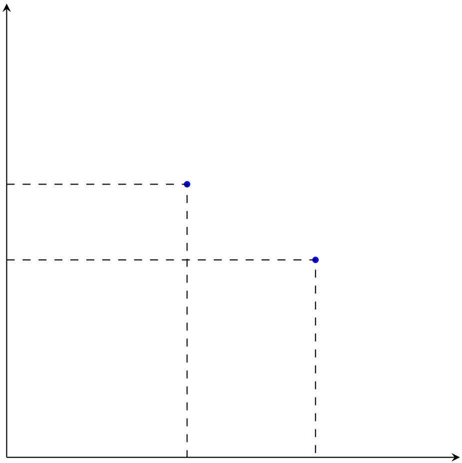
$K_1 = 0$



A graphical illustration for notations

$r_2 = 2$

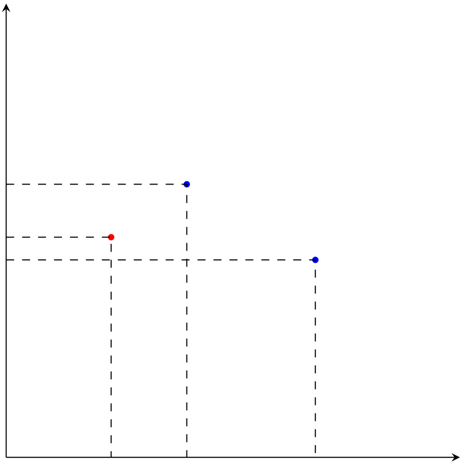
$K_2 = 0$



A graphical illustration for notations

$r_3 = 2$

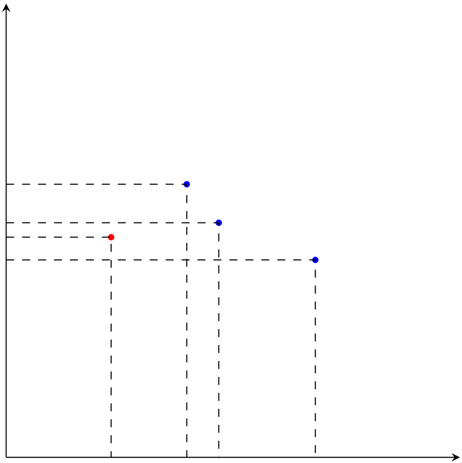
$K_3 = -1$



A graphical illustration for notations

$r_4 = 3$

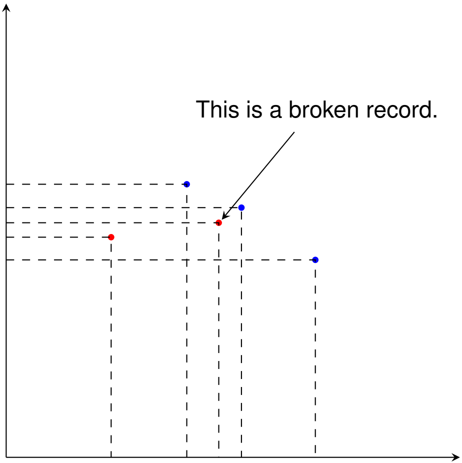
$K_4 = 0$



A graphical illustration for notations

$r_5 = 3$

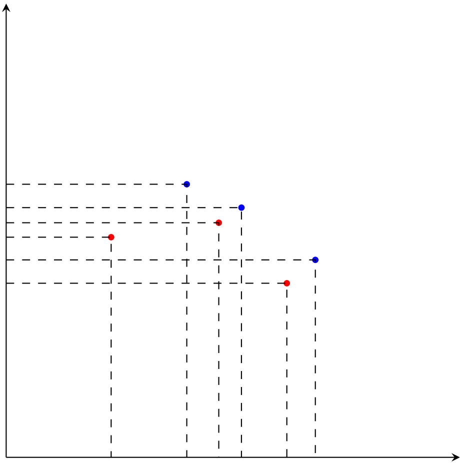
$K_5 = 1$



A graphical illustration for notations

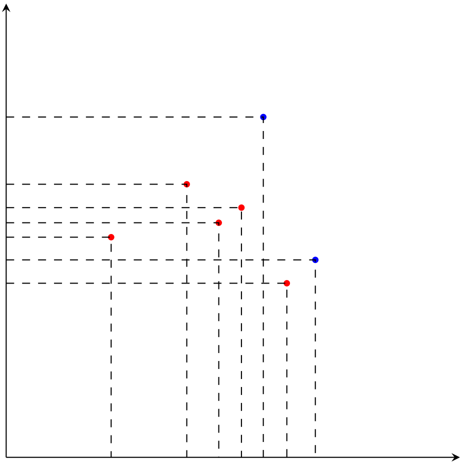
$r_6 = 3$

$K_6 = -1$



A graphical illustration for notations

$$r_7 = 2$$
$$K_7 = 2$$



Models of our interest

In the study of record-breaking, we are interested in two models where i.i.d. observations $(\mathbf{X}^{(n)})_{n\geq 1}$ are uniformly distributed in

1. the d-dimensional unit cube $\mathcal{C}_d := [0,1]^d$;
2. the d-dimensional unit simplex

$$\mathcal{S}_d := \{\mathbf{x} = (x_1, \ldots, x_d) \in \mathbb{R}^d : x_i \geq 0 \text{ for all } i \in [d] \text{ and } \|\mathbf{x}\|_{\ell^1} \leq 1\}.$$

Note that our first model covers the case where all the components of each observation are collectively independent, each may in fact have a *different* continuous distribution.

Our second model serves as a prototype for negatively associated coordinates in each observation.

Previous work

- cube case: exact and asymptotic expressions for $\mathbb{E}\, r_n(d)$, and those for $\operatorname{Var} r_n(3)$ ($\operatorname{Var} r_n(2)$ follows from the variance of Poisson-binomial sum). Barndorff-Nielsen and Sobel (1966)
- cube case: asymptotics for the variance of $r_n(d)$ for general d. Bai et al. (1998)
- asymptotic normality and Berry–Esseen bounds are proved for $r_n(2)$ associated with i.i.d. samples uniformly distributed in more general planar regions. Bai et al. (2001) and Bai, Hwang, and Tsai (2003)
- cube case: asymptotic normality for $r_n(d)$ for general d–a very sketchy proof. Baryshnikov (2000)
- cube case: asymptotic normality and Berry–Esseen bounds for $r_n(d)$ for general d by Stein's method; simplex case: claimed without a proof. Bai et al. (2005)

Previous work

- cube case: analysis of an importance sampling algorithm to generate Pareto records. Fill and Naiman (2019)
- cube case: frontiers for the record-setting region. Fill and Naiman (2020)
- cube case: limiting distribution for the conditional distribution of $K_n(2)$ given $K_n(2) \geq 0$. Fill (2021)
- cube case: limiting distribution for the conditional distribution of $K_n(d)$ given $K_n(d) \geq 0$ by the method of moments (conditionally). Fill (2023)

Table of Contents

Empirical experiment for observations in $\mathcal{C}_2$

The study of the record-breaking process $K(d)$ was motivated by the empirical data in the following table in Fill and Naiman (2019, Table 1):

k	N_k	$\tilde{p}_k$
0	50,334	0.50334
1	24,667	0.24667
2	12,507	0.12507
3	6,335	0.06335
4	3,040	0.03040
5	1,571	0.01571
...	...	...

Results of simulation experiment in which $100{,}000$ bivariate records are generated. The number of records that break k current records is denoted by N_k, and $\tilde{p}_k = N_k/100{,}000$ is the fraction of the $100{,}000$ records that break k records.

Empirical pattern and previous work

The data in that table reveal a $\mathrm{Geometric}(1/2)$ pattern for the proportion of the 100,000 bivariate records that break k records when i.i.d. samples are uniformly distributed in the unit square $[0, 1]^2$.

Empirical pattern and previous work

The data in that table reveal a $\mathrm{Geometric}(1/2)$ pattern for the proportion of the 100,000 bivariate records that break k records when i.i.d. samples are uniformly distributed in the unit square $[0,1]^2$.

Through analytical method, Fill (2021) proves that the conditional distribution of $K_n(2)$ given $K_n(2) \geq 0$ indeed converges weakly to $\mathrm{Geometric}(1/2)$.

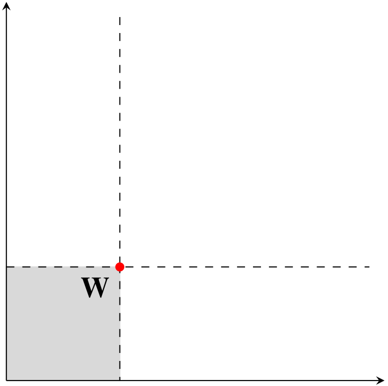
Fill (2021) only partially explains the pattern, as it doesn't reveal the limiting behavior of the fraction $\tilde{p}_k$ directly. See more in the section about ratio limit theorem.

Description of the limiting distribution for general d

What is the limit law of $\mathcal{L}(K_n(d) \mid K_n(d) \geq 0)$ for general d?

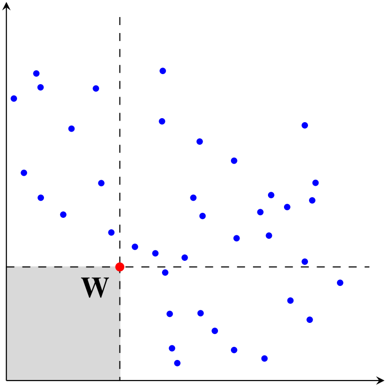
We can in fact compute the limit laws explicitly for both of our models when $d = 2$ (you have seen it for the cube case; the simplex case will be discussed later). As far as we know, there are no closed form expressions for them when $d \geq 3$. Instead, we describe these limit distributions as simple Exponential mixtures of distributions expressed in terms of Poisson point processes.The mixture is Exponential in the cube case, but not in the simplex case.

Description of the limiting distribution



$W \sim \text{Exponential}(1)$, set $\mathbf{W} = (W^{1/d}, \ldots, W^{1/d})$.

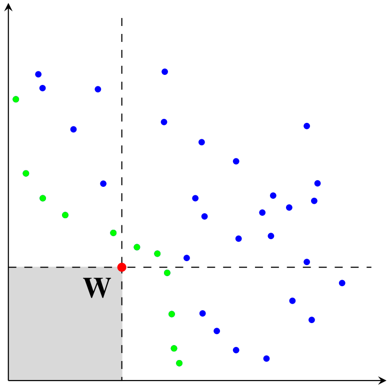
Description of the limiting distribution



A unit-rate homogeneous Poisson point process in the set

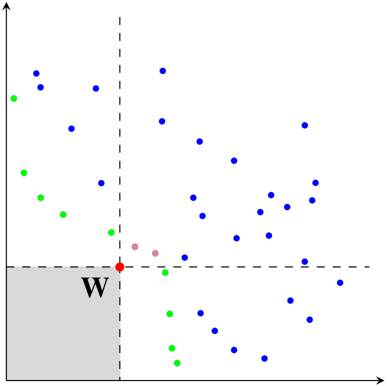
$$\{\mathbf{y} \in \mathbb{R}^d_{>0} : \mathbf{y} \not\prec \mathbf{W}\}.$$

Description of the limiting distribution



Find the minima from the Poisson point process.

Description of the limiting distribution



Count the number of minima that dominate $\mathbf{W}$.

Limiting distribution for the cube case

Cube Theorem [Fill (2023), S and Fill (2023+)]

Let $(\mathbf{X}^{(n)})_{n\geq 1}$ be a sequence of i.i.d. random vectors, each uniformly distributed in the unit cube $\mathcal{C}_d$. Then K_n, conditionally given $K_n \geq 0$, converges in distribution as $n \to \infty$ to the law of $\mathcal{K}_{\text{cube}} \equiv \mathcal{K}_{\text{cube}}(d)$, where

$$\mathcal{L}(\mathcal{K}_{\text{cube}}) \text{ is the mixture } \mathbb{E}\,\mathcal{L}(\mathcal{K}_W).$$

Here W has $\text{Exponential}(1)$ distribution, and for $\mathbf{w} = (w^{1/d}, \dots, w^{1/d})$

$$\mathcal{K}_w := \text{number of minima that dominate } \mathbf{w} \in \mathbb{R}^d_{>0}$$

in a unit-rate homogeneous Poisson point process (PPP) in the set

$$\{\mathbf{y} \in \mathbb{R}^d_{>0} : \mathbf{y} \not< \mathbf{w}\}.$$

Limiting distribution for the simplex case

Simplex Theorem [S and Fill (2023+)]

Let $(\mathbf{X}^{(n)})_{n \in \mathbb{N}}$ be a sequence of i.i.d. random vectors, each uniformly distributed in the unit simplex $\mathcal{S}_d$, then K_n, conditionally given $K_n \geq 0$, converges in distribution as $n \to \infty$ to the law of $\mathcal{K}_{\text{simplex}} \equiv \mathcal{K}_{\text{simplex}}(d)$, where

$$\mathcal{L}(\mathcal{K}_{\text{simplex}}) \text{ is the mixture } \mathbb{E}\,\mathcal{L}(\mathcal{K}'_V).$$

Here V has the same distribution as $Z^{1/d}$, where the random variable Z is distributed as $\text{Gamma}(1/d, 1)$, and for $\mathbf{v} = (v/d, \ldots, v/d)$

$$\mathcal{K}'_v := \text{number of minima that dominate } \mathbf{v} \in \mathbb{R}^d$$

in a homogeneous Poisson point process with rate $d!$ in the set

$$\{\mathbf{y} \in \mathbb{R}^d : \mathbf{y}_+ \geq 0, \mathbf{y} \not\prec \mathbf{v}\}.$$

Remark

The Cube Theorem has been proved in Fill (2023) largely analytically using the method of moments (conditionally). The treatment therein requires evaluating the higher-order moments of both the conditional distribution of K_n and its limiting distribution; this approach relies on intricate calculations and provides little insight as to why the particular weak limit in the Cube Theorem should emerge.

Remark

Our *probabilistic arguments* are quite robust. They can be adapted easily to prove the Simplex Theorem *without* too much additional effort. We also find that the Poisson point process appearing in the description of the limiting distribution is essentially a consequence of the so-called "law of rare events" or "law of small numbers", i.e., the Poisson approximation to the binomial distribution. We anticipate that probabilistic approach would make other extensions (e.g. sampling from Dirichlet distribution, etc.) fairly routine.

Notations

Definition

The *(open) positive orthant* translated by $\mathbf{x}$ is defined as the set

$$O_{\mathbf{x}}^{+} := \{\mathbf{y} \in \mathbb{R}^d : \mathbf{y} \succ \mathbf{x}\}.$$

The *(open) negative orthant* translated by $\mathbf{x}$ is defined as the set

$$O_{\mathbf{x}}^{-} := \{\mathbf{y} \in \mathbb{R}^d : \mathbf{y} \prec \mathbf{x}\}.$$

Heuristics: convergence of random measures

Let S be a locally compact Polish space, and $\mathscr{B}(S)$ be the Borel σ-algebra on S. Let λ be Lebesgue measure on $\mathbb{R}$.

Proposition [Resnick (2008, Proposition 3.21)]

For each $n \in \mathbb{N}$, suppose that $(X_{n,j})_{j\in\mathbb{N}}$ is a sequence of i.i.d. random variables taking values in $(S, \mathscr{B}(S))$. Let μ be a nonnegative σ-finite Radon measure on $\mathscr{B}(S)$. Define

$$N_n := \sum_{j=1}^{\infty} \delta_{(j/n, X_{n,j})},$$

and N is a Poisson random measure (PRM) on $\mathbb{R}_{\geq 0} \times S$ with mean measure $\lambda \otimes \mu$. Then N_n converges weakly to N if and only if

$$n\mathbb{P}(X_{n,1} \in \cdot) \xrightarrow{\text{w}} \mu \quad \text{on } S.$$

Unit cube: a heuristic argument for the Poisson point process

We consider record-*small* values (which we will call *small records*, for short) in this case.

Conditionally given $\mathbf{X}^{(n)} = \mathbf{x}^{(n)}$ and $K_n \geq 0$, the random vectors $\mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(n-1)}$ are i.i.d., their common distribution is uniform on $\mathcal{C}_d \setminus O^-_{\mathbf{x}^{(n)}}$.

Scaling each observation by a factor $n^{1/d}$, and suppose $n^{1/d}\mathbf{x}^{(n)} \to \mathbf{x}$ (We choose the factor $n^{1/d}$ for simplicity, in the actual proof, we scale each component of $\mathbf{x}^{(n)}$ by a different carefully chosen factor to make this true) as $n \to \infty$. Then, by previous proposition, the conditional distribution of the random measure

$$\sum_{i=1}^{n-1} \delta_{n^{1/d}\mathbf{X}^{(i)}}$$

converges weakly to the PRM on $\mathbb{R}^d_{>0} \setminus O^-_{\mathbf{x}}$ with mean measure λ.

Table of Contents

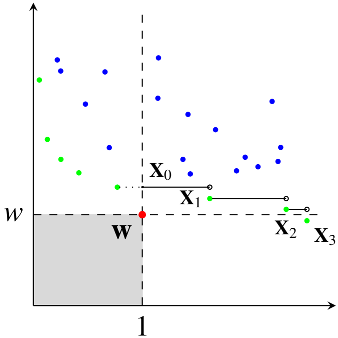
Unit square: bivariate records

In connection with the Cube Theorem, we identify a time-homogeneous Markov jump process hidden behind the Poisson point process in the limiting distribution. We then use this process to prove the main $\mathrm{Geometric}(1/2)$ result in Fill (2021) as a corollary to the Cube Theorem.

Corollary 1

Let $(\mathbf{X}^{(n)})_{n\geq 1}$ be a sequence of i.i.d. random vectors, each uniformly distributed in the unit square $\mathcal{C}_2$. Then the distribution of $\mathcal{K}_{\mathrm{cube}}(2)$ is $\mathrm{Geometric}(1/2)$ with support $\mathbb{Z}_{\geq 0}$.

An illustration of the Markov jump process



In this realization of the Poisson point process of the Cube Theorem, the minima are marked in green. The observations $\mathbf{X}_1$ and $\mathbf{X}_2$ are the two minima that dominate $\mathbf{w}$.

Unit simplex: bivariate records

One can transform minima in $\mathbb{R}^2$ to those in $\mathbb{R}$, an intensively studied subject, by employing concomitants. In this way, we identify the distribution of $\mathcal{K}_{\text{simplex}} \equiv \mathcal{K}_{\text{simplex}}(2)$ in terms of common distributions as a corollary to the Simplex Theorem.

The probability generating function of $\mathcal{K}_{\text{simplex}}$ is

$$\mathbb{E}\, s^{\mathcal{K}_{\text{simplex}}} = 2^{s-2} + \frac{1 - 2^{s-2}}{2 - s}.$$

Unit simplex: bivariate records

One can transform minima in $\mathbb{R}^2$ to those in $\mathbb{R}$, an intensively studied subject, by employing concomitants. In this way, we identify the distribution of $\mathcal{K}_{\text{simplex}} \equiv \mathcal{K}_{\text{simplex}}(2)$ in terms of common distributions as a corollary to the Simplex Theorem.

The probability generating function of $\mathcal{K}_{\text{simplex}}$ is

$$\mathbb{E}\, s^{\mathcal{K}_{\text{simplex}}} = 2^{s-2} + \frac{1 - 2^{s-2}}{2 - s}.$$

Corollary 2 [S and Fill (2023+)]

$$\mathbb{P}(\mathcal{K}_{\text{simplex}}(2) = k) = \frac{1}{4}\frac{(\ln 2)^k}{k!} + \frac{1}{2^{k+1}} - \frac{1}{2^{k+3}} \sum_{j=0}^{k} \frac{(2\ln 2)^j}{j!}.$$

Unit simplex: bivariate records

One can transform minima in $\mathbb{R}^2$ to those in $\mathbb{R}$, an intensively studied subject, by employing concomitants. In this way, we identify the distribution of $\mathcal{K}_{\text{simplex}} \equiv \mathcal{K}_{\text{simplex}}(2)$ in terms of common distributions as a corollary to the Simplex Theorem.

The probability generating function of $\mathcal{K}_{\text{simplex}}$ is

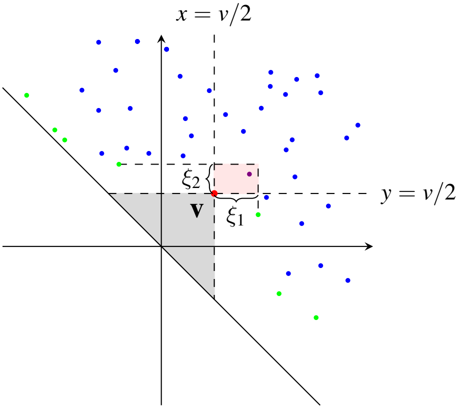
$$\mathbb{E}\, s^{\mathcal{K}_{\text{simplex}}} = 2^{s-2} + \frac{1-2^{s-2}}{2-s}.$$

Corollary 2 [S and Fill (2023+)]

$$\mathbb{P}(\mathcal{K}_{\text{simplex}}(2) = k) = \frac{1}{4}\frac{(\ln 2)^k}{k!} + \frac{1}{2^{k+1}} - \frac{1}{2^{k+3}} \sum_{j=0}^{k} \frac{(2\ln 2)^j}{j!}.$$

This term predominates asymptotically due to cancellations!

An illustration of the projection idea



In the Poisson point process of the Simplex Theorem above, all but one of the minima are marked in green. Exactly one of the minima (marked in violet) dominates the point $\mathbf{v}$ marked in red.

Table of Contents

From an unbounded region to a bounded one

Since generating Pareto records has a high computational cost (especially in high dimensions), in order to study the empirical behavior of the record-breaking process, one might be interested in perfectly simulating $\mathcal{L}(\mathcal{K}_{\text{cube}}(d))$ or $\mathcal{L}(\mathcal{K}_{\text{simplex}}(d))$.

From an unbounded region to a bounded one

Since generating Pareto records has a high computational cost (especially in high dimensions), in order to study the empirical behavior of the record-breaking process, one might be interested in perfectly simulating $\mathcal{L}(\mathcal{K}_{\text{cube}}(d))$ or $\mathcal{L}(\mathcal{K}_{\text{simplex}}(d))$.

Recall that $\mathcal{K}_{\text{cube}}(2)$ has the Geometric$(1/2)$ distribution; one can easily and efficiently simulate from this distribution. In light of Corollary 2, it is also elementary to simulate from the distribution of $\mathcal{K}_{\text{simplex}}(2)$.

From an unbounded region to a bounded one

Since generating Pareto records has a high computational cost (especially in high dimensions), in order to study the empirical behavior of the record-breaking process, one might be interested in perfectly simulating $\mathcal{L}(\mathcal{K}_{\text{cube}}(d))$ or $\mathcal{L}(\mathcal{K}_{\text{simplex}}(d))$.

Recall that $\mathcal{K}_{\text{cube}}(2)$ has the Geometric$(1/2)$ distribution; one can easily and efficiently simulate from this distribution. In light of Corollary 2, it is also elementary to simulate from the distribution of $\mathcal{K}_{\text{simplex}}(2)$.

For general d, the limiting distribution is described in terms of a Poisson point process in a unbounded region. We need to come up with a way to reduce it to sampling from a bounded region.

A perfect simulation algorithm: cube case

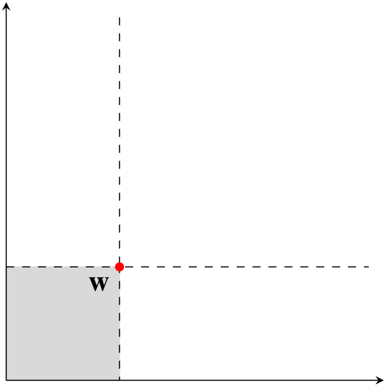
A routine for generating an observation from $\mathcal{L}(\mathcal{K}_{\text{cube}}(d))$ is described as follows. A similar routine works for generating observations from $\mathcal{L}(\mathcal{K}_{\text{simplex}}(d))$.

- Generate a random variable $W = w$ from the $\text{Exponential}(1)$ distribution.
- Independently generate d independent $\text{Exponential}(w^{(d-1)/d})$ random variables $\xi_1 = x_1, \ldots, \xi_d = x_d$.
- Independently simulate a unit-rate Poisson process in the region

$$\prod_{i=1}^{d}[0, w^{1/d} + x_i] \setminus \prod_{i=1}^{d}[0, w^{1/d}].$$

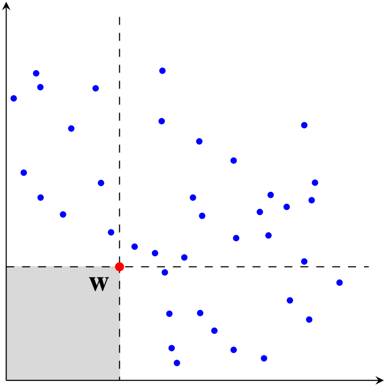
- Determine and record the number of minima from this Poisson process that dominate $(w^{1/d}, \cdots, w^{1/d})$ as a sample from $\mathcal{L}(\mathcal{K}_{\text{cube}}(d))$.

An illustration for the perfect simulation algorithm in $\mathbb{R}^2$



Generate $W = w$ from $\mathrm{Exponential}(1)$ and set $\mathbf{w} = (w^{1/2}, w^{1/2})$.

An illustration for the perfect simulation algorithm in $\mathbb{R}^2$

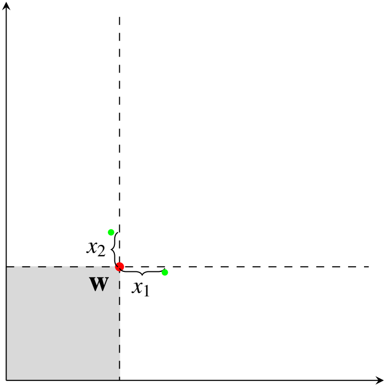


Original: a unit-rate homogeneous Poisson point process in the set

$$\{\mathbf{y} \in \mathbb{R}^d_{>0} : \mathbf{y} \not< \mathbf{w}\}$$

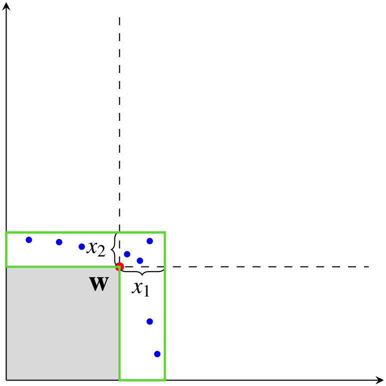
.

An illustration for the perfect simulation algorithm in $\mathbb{R}^2$



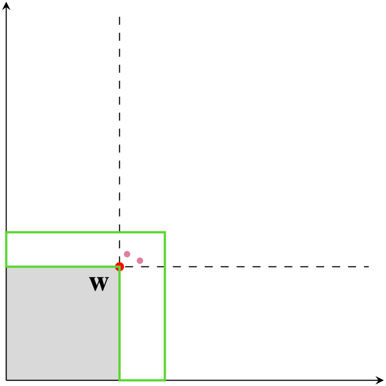
Now: independently generate two independent $\text{Exponential}(w^{1/2})$ random variables $\xi_1 = x_1$ and $\xi_2 = x_2$.

An illustration for the perfect simulation algorithm in $\mathbb{R}^2$



Independently simulate a unit-rate Poisson process in the region enclosed in green.

An illustration for the perfect simulation algorithm in $\mathbb{R}^2$



Count the number of minima from the process that dominate $\mathbf{w}$.

Thank you!

References I

Arnold, B.C., N. Balakrishnan, and H.N. Nagaraja (1998). *Records*. Wiley Series in Probability and Statistics. New York: Wiley. ISBN: 9781118150405.

Bai, Zhi-Dong, Hsien-Kuei Hwang, and Tsung-Hsi Tsai (2003). "Berry-Esseen bounds for the number of maxima in planar regions". In: *Electron. J. Probab.* 8.9, pp. 1–26. ISSN: 1083-6489. DOI: 10.1214/EJP.v8-137. URL: https://doi.org/10.1214/EJP.v8-137.

Bai, Zhi-Dong et al. (1998). "On the variance of the number of maxima in random vectors and its applications". In: *Annals of Applied Probability* 8.3, pp. 886–895. DOI: 10.1214/aoap/1028903455. URL: https://doi.org/10.1214/aoap/1028903455.

References II

Bai, Zhi-Dong et al. (2001). "Limit theorems for the number of maxima in random samples from planar regions". In: *Electron. J. Probab.* 6.3, pp. 1–41. ISSN: 1083-6489. DOI: 10.1214/EJP.v6-76. URL: https://doi.org/10.1214/EJP.v6-76.

Bai, Zhi-Dong et al. (2005). "Maxima in hypercubes". In: *Random Structures & Algorithms* 27.3, pp. 290–309. DOI: https://doi.org/10.1002/rsa.20053. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/rsa.20053. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/rsa.20053.

References III

- Barndorff-Nielsen, O. and M. Sobel (1966). "On the Distribution of the Number of Admissible Points in a Vector Random Sample". In: *Theory of Probability & Its Applications* 11.2, pp. 249–269. DOI: 10.1137/1111020. URL: https://doi.org/10.1137/1111020.
- Baryshnikov, Yu. (2000). "Supporting-points processes and some of their applications". In: *Probab. Theory Related Fields* 117.2, pp. 163–182. ISSN: 0178-8051. DOI: 10.1007/s004400050002. URL: https://doi.org/10.1007/s004400050002.
- Fill, James Allen (2021). "Breaking bivariate records". In: *Combinatorics, Probability and Computing* 30.1, pp. 105–123. DOI: 10.1017/S0963548320000309.

- Fill, James Allen (2023). "Breaking multivariate records". In: *Electronic Journal of Probability* 28.none, pp. 1 –27. DOI: 10.1214/23-EJP968. URL: https://doi.org/10.1214/23-EJP968.
- Fill, James Allen and Daniel Q. Naiman (2019). *Generating Pareto records*. arXiv: 1901.05621.
- — (2020). "The Pareto record frontier". In: *Electronic Journal of Probability* 25.92, pp. 1–24. DOI: 10.1214/20-EJP492. URL: https://doi.org/10.1214/20-EJP492.

References V

Resnick, Sidney I. (2008). *Extreme values, regular variation and point processes*. Springer Series in Operations Research and Financial Engineering. Reprint of the 1987 original. Springer, New York, pp. xiv+320. ISBN: 978-0-387-75952-4.