Electronic Journal of Probability

Vol. 12 (2007), Paper no. 26, pages 767–786.

Journal URL
http://www.math.washington.edu/~ejpecp/

The Chain Records

Alexander V. Gnedin
Department of Mathematics, Utrecht University
Postbus 80010, 3508 TA Utrecht, The Netherlands
E-mail address: gnedin@math.uu.nl

Abstract

Chain records is a new type of multidimensional record. We discuss how often the chain records occur when the background sampling is from the unit cube with uniform distribution (or, more generally, from an arbitrary continuous product distribution in d dimensions). Extensions are given for sampling from more general spaces with a self-similarity property.

Key words: multidimensional records, random orders, chains, Ewens partition.

AMS 2000 Subject Classification: Primary 60G70, 60C05.

Submitted to EJP on August 28, 2006, final version accepted May 4, 2007.

1 Introduction

Consider independent marks $X_1, X_2, \dots$ sampled from the uniform distribution in $Q_d = [0,1]^d$. We define a mark X_n to be a *chain record* if X_n breaks the last chain record in $X_1, \dots, X_{n-1}$. More precisely, *record values* and *record indices* are introduced recursively, by setting $T_1 = 1$, $R_1 = X_1$ and

$$T_k = \min\{n > T_{k-1} : \ X_n \prec R_{k-1}\}\,, \quad R_k = X_{T_k}\,, \qquad k > 1\,.$$

Here, $\prec$ denotes the standard strict partial order on $\mathbb{R}^d$ defined in terms of component-wise orders by

$$x = (x^{(1)}, \dots, x^{(d)}) \prec y = (y^{(1)}, \dots, y^{(d)}) \ \text{ iff } \ x \neq y \ \text{ and } \ x^{(i)} \leq y^{(i)} \ \text{ for } \ i = 1, \dots, d\,.$$

It is easy to see that, in any dimension d, the terms of (T_k) are indeed well defined for all k, that is the chain records occur infinitely many times.

The term *chain record* is chosen to stress the most straightforward 'greedy' method of constructing a chain in a random partial order. This notion of multidimensional record has not been explored so far, although the definition is an obvious restatement of the classical recursive definition of a strict lower record (1; 37). The point advocated here is that the chain records are closer relatives to the classical univariate records than various other types of multidimensional records studied in the literature (4; 5; 13; 19; 20; 21; 26; 30; 31; 32; 34; 35). Apart from the theory of extreme values, the chain records are of interest in some other contexts, like search trees (14; 17), partially exchangeable partitions (38) and fragmentation processes (7). We shall exploit these connections to analyse the frequency of chain records and related properties of their values.

The chain records are intermediate between two other types of multidimensional records. We say that a *strong record* occurs at index n if either $n = 1$, or $n > 1$ and

$$X_n \prec X_j \ \text{ for } \ j = 1, \dots, n-1.$$

In the terminology of partially ordered sets, a strong record X_n is the *least* element in the point set $\{X_1, \dots, X_n\}$. Since repetitions in each component have probability zero, X_n is a strong record if and only if there is a strict lower record in each of d components, that is there are d *marginal* records simultaneously. We say that a *weak record* occurs at index n if either $n = 1$, or $n > 1$ and

$$X_j \not\prec X_n \ \text{ for } \ j = 1, \dots, n-1.$$

A weak record X_n is a *minimal* element in the set $\{X_1, \dots, X_n\}$. Obviously, each strong record is a chain record. Also, each chain record is a weak record, as it follows easily by induction from transitivilty of the relation $\prec$. Thus, denoting $\underline{N}_n, \overline{N}_n$ and N_n the counts of strong, weak and chain records among the first n marks, respectively, we have

$$\underline{N}_n \leq N_n \leq \overline{N}_n\,.$$

To illustrate, for the two-dimensional configuration of points in Figure 1 the weak records occur at times 1, 2, 3, 5, 6, 7, 8, 9, the strong records occur at times 1, 8, the marginal records occur at times $1, 2, 3, 5, 6$, and the chain records occur at times 1, 5, 8.

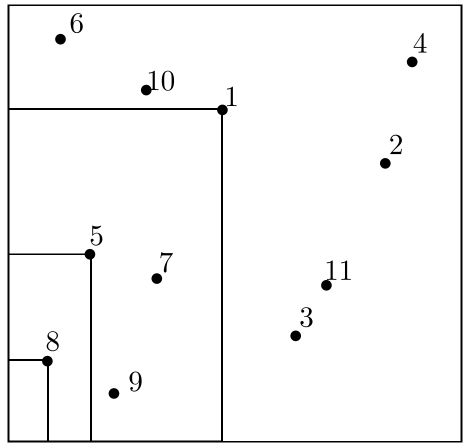

Figure 1: Chain records in the square

Clearly, all three types of records coincide in dimension one, but for $d > 1$ they are very different. For instance, unlike two other types, the chain records are sensitive to arrangement of marks in the sequence: a permutation of $X_1, \ldots, X_{n-1}$ may destroy or create a chain record at index n. To show quantitative differences we shall compare how often the records of different kinds may occur.

Recall that in dimension $d = 1$ the question about the frequency of records is settled by the Dwass-Rényi lemma (8, Lemma 5.1), which states that the indicators of records $I_n = 1(N_n > N_{n-1})$ are independent, with probability $p_n = \mathbb{P}(I_n = 1) = 1/n$ for a record at index n. This basic fact can be concluded by combinatorial arguments from exchangeability and uniqueness of the minimum among n sample points. It follows then in a standard way that N_n is asymptotically Gaussian with both mean and variance asymptotic to $\log n$.

Properties of the strong-record counts for sampling from Q_d are also rather simple. By independence of the marginal rankings we have a representation $\underline{N}_n = \underline{I}_1 + \ldots + \underline{I}_n$ with independent Bernoulli indicators and $\underline{p}_n := \mathbb{P}(\underline{I}_n = 1) = n^{-d}$. Thus

$$\mathbb{E}\,\underline{N}_n = \sum_{j=1}^{n} \frac{1}{j^d}\,.$$

For $d > 1$ the series $\sum \underline{p}_n$ converges, hence the total number of the strong records is finite with probability one. This implies that $\underline{N}_n$ converge almost surely to some random variable which is not Gaussian.

Counting the weak records is a more delicate matter since the indicators $\overline{I}_n = 1(\overline{N}_n > \overline{N}_{n-1})$ are not independent. However, there is a simple device to reduce counting the weak records to known results about the minimal points. To this end, let $\xi_1 < \ldots < \xi_n$ be n uniform order statistics, independent of the X_j's. Observe that the sequence $(X_1, \xi_1), \ldots, (X_n, \xi_n)$ is distributed like a uniform sample of n points from Q_{d+1}, arranged by increase of the last coordinate. A minute thought shows that X_j $(j \le n)$ is a weak record if and only if (X_j, ξ_j) is a minimal point among

n points in $d+1$ dimensions. Using this correspondence and (4, Equation (3.39)) we have

$$\mathbb{E}\left[\overline{N}_n\right] = \sum_{1\le j_1\le \ldots \le j_d \le n} \frac{1}{j_1\cdots j_d} \sim \frac{1}{d!}\left(\log n\right)^d ,$$

(by exchangeability, $n\overline{p}_n = \mathbb{E}\left[\overline{N}_n\right]$). Similarly, from the results in (5) we can conclude that the variance $\mathrm{Var}\left[\overline{N}_n\right]$ is of the same order $(\log n)^d$, and that $\overline{N}_n$ is asymptotically Gaussian. Asymptotic expansions of the variance of the number of minimal points are given in (3).

Thus the strong records are much more rare and the weak records are much more frequent than the classical records. In this paper we show that, as far as the frequency is concerned, the chain records in any dimension d are more in line with the classical records:

Proposition 1. *For sampling from Q_d with uniform distribution the number of chain records N_n is approximately Gaussian with moments*

$$\mathbb{E}\left[N_n\right] \sim d^{-1}\log n\,, \quad \mathrm{Var}\left[N_n\right] \sim d^{-2}\log n\,.$$

The CLT will be proved in Section 3. Above that, we will derive exact and asymptotic formulas for the probability of a chain record $p_n = \mathbb{P}(N_n > N_{n-1})$ and will discuss some point-process scaling limits, generalizing that in the one-dimensional case.

The counting results remain the same if sampling from the uniform distribution in Q_d is replaced by sampling from any continuous product distribution in $\mathbb{R}^d$. The situation changes radically for sampling from more general probability measures μ on $\mathbb{R}^d$, since the properties of record counts are no longer 'distribution free'. For instance, $\underline{p}_n \sim cn^{-2/(1+\rho)}(\log n)^{-\rho/\rho+1}$ for the bivariate correlated Gaussian distribution (26, p. 13) (in which case the total number of strong records is finite), while $\underline{p}_n \sim c/n$ for the bivariate Cauchy distribution (20, p. 515) (in which case the total number of strong records is infinite). The distribution of $\overline{N}_n$ for various classes of measures μ was studied e.g. in (5; 6; 20; 33).

More generally, one can define records for sampling from some probability space $(\mathcal{X}, \mathcal{F}, \mu)$ endowed with a strict partial order $\prec$. In these general terms it is easy to give a criterion for there to be infinitely many chain records. Also, Proposition 1 and scaling limit results generalize for a class of sampling spaces which possess a self-similarity property, versions of which were exploited in a number of related contexts (12; 14; 24).

2 The heights at chain records and stick-breaking

For $x \in Q_d$ the quadrant $L_x := \{y \in Q_d : y \prec x\}$ is the lower section of the partial order at x. The *height* $h(x)$ is the measure of L_x, which in the case of uniform distribution under focus is equal to the product of coordinates. The height is a key quantity to look at, because the heights at chain records determine the *sojourns* $T_{k+1} - T_k$ (which may be also called inter-record times). Let $H_k = h(R_k)$ be the height at the kth chain record. Because a new chain record R_{k+1} is established as soon as L_{R_k} is hit by some mark, we have, exactly as in the classical case (8, Theorem 4.1):

Lemma 2. *Given the sequence (H_k) the sojourns $T_{k+1} - T_k$ are conditionally independent, geometrically distributed with parameters H_k, $k = 1, 2, \ldots$*

Note that we can express the record counts and indicators through (T_k) as

$$N_n = \#\{k : T_k \le n\}, \quad I_n = 1(n \in \{T_k;\, k = 1, 2, \ldots\}), \tag{1}$$

which implies that the occurences of the chain records are completely determined by the law of (H_k). Note also that $N_n \ge k$ is equivalent to $T_k \le n$. The lemma has the following elementary but important consequence.

Corollary 3. *Given (H_k), the conditional law of (I_n) is the same as for the classical records from Q_1 with the uniform distribution.*

The heights at chain records undergo a multiplicative renewal process, sometimes called *stick-breaking*. Define random variables $W_1, W_2, W_3, \ldots$ by setting

$$H_k = W_1 \cdots W_k\,, \qquad k = 1, 2, \ldots \tag{2}$$

Lemma 4. *The random variables $W_1, W_2, W_3, \ldots$ are independent replicas of a random variable $W := H_1$.*

Proof. Each lower section L_x, viewed as a partially ordered probability space with normalized Lebesgue measure is isomorphic to Q_d (via a coordinate-wise scale transformation). Hence all ratios H_{k+1}/H_k are i.i.d. □

Explicitly, the density of W is

$$\mathbb{P}(W \in \mathrm{d}s) = \frac{|\log s|^{d-1}}{(d-1)!}\,\mathrm{d}s\,, \quad s \in [0,1], \tag{3}$$

and the Mellin transform is

$$g(\lambda) := \mathbb{E}\,[W^\lambda] = (\lambda + 1)^{-d}\,. \tag{4}$$

as follows by noting that H_1 is the product of d independent uniform variables.

It is an easy exercise to check for small sample sizes that the I_n's are not independent for $d > 1$, in contrast to the case $d = 1$. Thus for W with uniform distribution the independence holds, and for W with density (3) ($d > 1$) not. This motivates a more general question of characterizing all distributions for W which force independence of the indicators. Formally, let (H_k) be a stick-breaking sequence (2), with W_j's being independent copies of an arbitrary variable W assuming values in $]0, 1[$. Let (T_k) be conditionally independent and geometric, as in Lemma 2, and define indicators by (1).

Proposition 5. *The indicators $I_1, I_2, \ldots$ are independent if and only if for some $\theta > 0$ the variable W has* beta$(\theta, 1)$ *density*

$$\mathbb{P}(W \in \mathrm{d}s) = \theta s^{\theta - 1}\mathrm{d}s\,, \qquad 0 < s < 1.$$

In this case $\mathbb{P}(I_n = 1) = \theta/(\theta + n - 1)$.

Proof. Let $v_n = \mathbb{P}(I_n = 0),\ u_n = 1 - v_n$. The 'if' part follows from the well-known sequential description of the Ewens partition (39).

To prove the 'only if' part start by observing that by Lemma 2 and from the assumption of independence of the indicators

$$\mathbb{P}((I_1, \ldots, I_n) = (1, 0, \ldots, 0)) = \mathbb{E}[(1 - H_1)^{n-1}] = v_2 \cdots v_n, \qquad n \geq 1, \qquad (5)$$

$$\mathbb{P}((I_1, \ldots, I_n) = (1, 0, \ldots, 0, 1, 1)) = \mathbb{E}[(1 - H_1)^{n-3} H_1 H_2] = v_2 \cdots v_{n-1} u_{n-1} u_n, \quad n \geq 3. \qquad (6)$$

By independence of $W_1 = H_1,\ W_2 = H_2/H_1$ and $W_1 =_d W_2$ the middle expression in (6) can be factored as

$$\mathbb{E}\left[(1 - H_1)^{n-3} H_1^2 \frac{H_2}{H_1}\right] = \mathbb{E}\left[(1 - H_1)^{n-3} H_1^2\right]\ \mathbb{E}[H_1].$$

This is expanded using $H_1^2 = 1 - 2(1 - H_1)^2 + (1 - H_1)^2$ and $H_1 = 1 - (1 - H_1)$ and evaluated by (5). Upon simplification (6) yields the identity

$$u_2 u_{n-1} = u_n(u_2 + u_{n-1} - u_2 u_{n-1}), \qquad n \geq 3.$$

For u_2 strictly between 0 and 1, the factor at u_n is positive thus $(u_n,\ n \geq 2)$ is uniquely determined by u_2. Setting $u_2 = \theta/(\theta + 1)$ the recursion is solved by $u_n = \theta/(\theta + n - 1)$. □

3 Random partitions and the CLT

The cube Q_d is decomposed in disjoint layers $Q \setminus L_{R_1}, L_{R_1} \setminus L_{R_2}, \ldots$. We define an ordered partition Π of the set $\mathbb{N}$ in disjoint nonempty blocks by assigning two integers i, j to the same block of the partition if and only if X_i, X_j fall in the same layer. The blocks are ordered by increase of their minimal elements, which are $T_1, T_2, \ldots$. For instance, for the configuration of points in Figure 1 the partition restricted to the set of 10 integers has three blocks $\{1, 2, 3, 4, 6, 10\}, \{5, 7, 9\}, \{8\}$. The partition Π is *partially exchangeable*, as introduced in (38), which means that for each n the probability of each particular value of $\Pi|_{[n]}$ depends only on the sequence of sizes of blocks of $\Pi|_{[n]}$, where $[n] := \{1, \ldots, n\}$. The sequence of sizes can be arbitrary composition (ordered partition) of integer n (it is $(6, 3, 1)$ in the above example). The number of chain records N_n is equal to the number of blocks of $\Pi|_{[n]}$.

More generally, a partition Π can be derived from arbitrary decreasing sequence of random variables (H_k) with values in $]0, 1[$, as suggested by Corollary 3. Let (U_j) be a i.i.d. sequence of $[0, 1]$ uniform points, independent of (H_k). We shall produce a transformed sequence $(U_j) \,|\, (H_k)$ by replacing some of the terms in (U_j) by the H_k's. Replace U_1 by H_1. Do not alter $U_2, U_3, \ldots$ as long as they do not hit $[0, H_1[$; then replace the first uniform point hitting the interval $[0, H_1[$ by H_2. Inductively, as $H_1, \ldots, H_k$ got inserted, keep on screening uniforms until first hitting $[0, H_k[$, then insert H_{k+1} in place of the uniform point that caused the hit, and so on. Eventually all H_k's will enter the resulting sequence. It is easy to see that given (H_k) the distribution of $(U_j) \,|\, (H_k)$ is the same as the conditional distribution of (U_j) given the subsequence of its record values (H_k). Now, the interval $]0, 1]$ is broken into subintervals $]H_{k+1}, H_k]$ (with $H_0 := 1$), and the partition Π is defined by assigning integers a, b to the same block if and only if the ath and the bth member of the sequence $(U_j) \,|\, (H_k)$ belong to the same subinterval.

In the classical case, (H_k) is the stick-breaking sequence with uniform factor W, and we have $(U_j)\,|\,(H_k) =_d (U_j)$, so the insertion does not alter the law of the uniform sequence. In this case Π is the *exhangeable* partition whose law corresponds to the Ewens sampling formula with parameter $\theta = 1$. If we take stick breaking with the beta factor W as in Proposition 5 we obtain the general Ewens partiton, which can be characterized by the property of independent indicators (36) (up to two trivial cases obtained by sending θ to 0 or ∞, respectively).

The construction of $(U_j)\,|\,(H_k)$ and Π does not impose any constraints on the law of the sequence (H_k), which can be an arbitrary nonincreasing sequence (the induced Π is then the most general partially exchangeable partition (38)). With this in mind, we shall take a more abstract approach and digress from the geometric description of records. Suppose (H_k) is defined by stick-breaking (2), with a generic random factor W assuming values in $]0,1[$. Let (T_k) be as in Lemma 2 and define N_n by (1). As above, T_k's and N_n can be interpreted in terms of a random partition Π.

Proposition 6. *Suppose W has finite logarithmic moments*

$$m = \mathbb{E}\,[-\log W], \quad \sigma^2 = \mathrm{Var}\,[-\log W].$$

Then for $n \to \infty$ the variable N_n is asymptotically Gaussian with moments

$$\mathbb{E}\,[N_n] \sim \frac{1}{m}\,\log n\,, \qquad \mathrm{Var}\,[N_n] \sim \frac{\sigma^2}{m^3}\,\log n\,.$$

We also have the strong law

$$N_n \sim \frac{1}{m}\,\log n \qquad \text{a.s.}$$

for which only $m < \infty$ is required.

Proof. Our strategy is to show that N_n is close to $K_n := \max\{k : H_k > 1/n\}$. By the renewal theorem (16) K_n is asymptotically Gaussian with the mean $m^{-1}\log n$ and the variance $\sigma^2 m^{-3}\log n$ because K_n is just the number of epochs on $[0, \log n]$ of the renewal process with steps $-\log W_j$.

By the construction of $(U_j)\,|\,(H_k)$, we have a dichotomy: $U_n \in\,]H_k, H_{k-1}]$ implies that either U_n will enter the transformed sequence or will get replaced by some $H_i \geq H_k$. Let $U_{n1} < \ldots < U_{nn}$ be the order statistics of $U_1, \ldots, U_n$. It follows that

(i) if $U_{nj} > H_k$ then $N_n \leq k + j$,

(ii) if $U_{nk} < H_k$ then $N_n \geq k$.

Let ξ_n be the number of uniform order statistics smaller than $1/n$. By definition, $H_{K_n+1} < 1/n < H_{K_n}$, hence K_n and ξ_n are independent and ξ_n is binomial$(n, 1/n)$. By (i), we have $N_n \leq K_n + \xi_n$ where ξ_n is approximately Poisson(1), which yields the desired upper bound.

Now consider the threshold $s_n = (\log n)^2/n$ and let $J_n := \max\{k : H_k > s_n\}$. By (ii), if the number of order statistics smaller than s_n is at least J_n then $N_n \geq J_n$. Because $\log n \sim \log n - 2\log\log n$ the index J_n is still asymptotically Gaussian with the same moments as K_n. On the other hand, the number of order statistics smaller than s_n is asymptotically Gaussian with moments about $(\log n)^2$. Hence elementary large deviation bounds imply that $N_n \geq J_n$ with probability very close to one. This yields a suitable lower bound, hence the CLT. Along the same lines, the strong law of large numbers follows from $N_n \sim K_n$. $\square$

Similar limit theorems have been proved by other methods for the number of blocks of exchangeable partitions (25), and for the length of a size-biased path in search trees (14; 18).

Proposition 1 follows as an instance of Proposition 6 by computing the logarithmic moments of the density (3)

$$m = \mathbb{E}\left[-\log W\right] = -g'(0) = d\,, \quad \sigma^2 = \mathrm{Var}\left[-\log W\right] = g''(0) - g'(0)^2 = d\,.$$

4 Poisson-paced records

The probability p_n of a chain record at index n is equal to the mean height of the last chain record before n. In terms of the quad-tree, p_n is the probability that X_n belongs to the path in the direction of the negative quadrant. To compute p_n we shall exploit the same kind of a continuous time process as in (24), which may be interpreted as the process of a tagged particle in a fragmentation with self-similarity index 1 (7). See (14; 15; 17; 18) for other approaches to functionals of quad-trees.

Let (τ_n) be the increasing sequence of points of a homogeneous Poisson point process (PPP) on $\mathbb{R}_+$, independent of the marks (X_n). The sequence $((X_n, \tau_n),\ n = 1, 2, \ldots)$ is then the sequence of points of a homogeneous PPP in $Q_d \times \mathbb{R}_+$ in the order of increase of the time component, which now assumes values in the continuous range $\mathbb{R}_+$. Let $\widehat{N}_t$ be the number of chain records and B_t the height of the last chain record on $[0, t]$, that is

$$\widehat{N}_t = \max\{k : \tau_{T_k} < t\}\,, \quad B_t = H_{\widehat{N}_t}.$$

Clearly, (B_t) is the predictable compensator for $(\widehat{N}_t)$, in particular

$$\mathbb{E}\left[\int_0^t B_s \mathrm{d}s\right] = \mathbb{E}\,[\widehat{N}_t]\,.$$

The effect of poissonization on the process of records amounts to just replacing geometric sojourns in Lemma 2 by exponentials. We again digress from the detailed geometric description, and construct a process (B_t) by first letting (H_k) to be a sequence of *distinct* visited states and then requiring that the sojourns in these states be conditionally independent, exponential(H_k). Assuming further that (H_k) is derived by stick-breaking (2) the process (B_t) is Markov time-homogeneous with a very simple type of behaviour. Given $B_t = b$ the process remains in state b for some rate-b exponential time and then jumps to a new state bW, with W a stereotypical random factor with values in $]0, 1[$.

Immediate from this description is the following self-similarity property: the law of (B_t) with initial state $B_0 = b$ is the same as the law of the process (bB_{bt}) with $B_0 = 1$. In this form the process is well defined for arbitrary initial state $b > 0$ and arbitrary W with values in the open interval. See (23) for features of this process related to the classical records and (7) for more general self-similar (also called semi-stable) processes derived from subordinators. Proposition 6 translates literally as a CLT for the number of jumps of (B_t) within a large time interval (in the language of (7), this is the number of dislocations of a tagged particle in a self-similar fragmentation process with finite dislocation measure).

By the self-similarity of (B_t) the moments

$$m_k(t) := \mathbb{E}\,[B_t^k]$$

satisfy a renewal-type equation

$$m_k'(t) = -m_k(t) + \mathbb{E}\,[W^k m_k(tW)]\,.$$

The series solution to this equation with the initial value $m_k(0) = 1$ is

$$m_k(t) = \sum_{k=0}^{\infty} \frac{(-t)^k}{k!} \prod_{j=0}^{k-1} (1 - g(j+k)), \tag{7}$$

with $g(\lambda) = \mathbb{E}[W^\lambda]$, as one can check by direct substitution (see (10)).
By (9, Theorem 1) the random variables tB_t converge, as $t \to \infty$, in distribution and with all moments to a random variable Y whose moments are given by

$$\mathbb{E}\,[Y^k] = \frac{1}{m} \prod_{r=1}^{k-1} \frac{r}{1 - g(r)}\,, \quad k = 1, 2, \ldots \tag{8}$$

where $m = -g'(0) = \mathbb{E}[-\log W]$. The law of Y is determined uniquely by the moments (8). This variable has a 'perpetuity series' representation

$$Y =_d E_0 W_0 + \sum_{k=1}^{\infty} E_k \prod_{j=0}^{k} W_j$$

where E_k's are exponential, W_j's for $j > 0$ are as before, W_0 has density

$$\mathbb{P}(W_0 \in \mathrm{d}s) = \frac{\mathbb{P}(W \le s)}{ms}\,\mathrm{d}s\,, \quad s \in [0,1] \tag{9}$$

(which is the stationary density for the stick-breaking process with factor W) and all variables are independent. Equivalently, we may write Y in the form of an exponential functional (11)

$$Y = \int_0^\infty e^{-S_t}\,\mathrm{d}t$$

for (S_t) a compound Poisson process with initial state $S_0 = -\log W_0$ and the jumps distributed like $-\log W$.
Connecting discrete and continuous time models we have

$$m_1(t) = e^{-t} \sum_{n=0}^{\infty} \frac{t^n}{n!}\,p_{n+1}\,,$$

which in terms of the chain records is the poissonization identity saying that $m_1(t)$ is the probability that the first arrival after t is a chain record. This implies, upon equating coefficients of the series,

$$p_n = \sum_{k=0}^{n-1} \binom{n-1}{k} (-1)^k \prod_{j=0}^{k-1} (1 - g(j+1))\,. \tag{10}$$

We remind that (10) with general $g(\lambda) = \mathbb{E}[W^\lambda]$ yields the probability that n is the minimal element in some block of Π, for Π a partition of $\mathbb{N}$ introduced in Section 3,

Specializing to the chain records from Q_d we substitute $g(\lambda) = (\lambda+1)^{-d}$ in the above formulas. Factoring

$$1 - g(j+k) = \prod_{r=1}^{d} \frac{j+k+1-e^{2\pi \mathrm{i} r/d}}{j+k+1}$$

we see that the series (7) is a generalized hypergeometric function of the type ${}_dF_d$. Formula (8) becomes

$$\mathbb{E}[Y^k] = \frac{(k!)^{d+1}}{kd} \prod_{r=2}^{k} \frac{1}{r^d - 1}. \tag{11}$$

The law of Y may be considered as a kind of extreme-value distribution, because it is the limit distribution of the height at the chain record last before time n. In the case $d = 1$ we recover well-known $Y =_d E$ with E standard exponential, and for $d = 2$ we get $Y =_d EU$ with E and U independent exponential and uniform random variables.

For $d = 1$ we obtain from (10) the familiar $p_n = 1/n$, and for $d = 2$ we obtain surprisingly simple expression $p_n = 1/(2n)$ (for $n > 1$). For $d > 2$ the formulas for p_n do not simplify. The depoissonization of the $k = 1$ instance implies, quite expectedly,

$$p_n \sim \frac{1}{d\,n}, \quad \text{as } \ n \to \infty\,.$$

Formula (10) can be compared with the analogous formulas

$$\underline{p}_n = g(n-1), \qquad \overline{p}_n = \sum_{k=0}^{n-1} \binom{n-1}{k} (-1)^k g(k)$$

for the occurences of strong and weak records. The last two follow by elementary combinatorics, but we do not have a direct combinatorial argument for (10).

5 Scaling limits

Let $b > 0$ be a scaling parameter which we will send to ∞. For $\{(R_k, T_k);\ k = 1, 2, \ldots\}$ the bivariate point process of record values and record times induced by sampling from Q_1, the scaled process $\{(bR_k,\, T_k/b), k = 1, 2, \ldots\}$ has a weak limit, which can be identified with the random set of minimal points of a homogeneous PPP in $\mathbb{R}_+^2$. Marginally, $\{bR_k\}$ and $\{T_k/b\}$ both converge to a scale-invariant PPP on $\mathbb{R}_+$ with intensity $\mathrm{d}x/x$. See (2; 29) on properties of the scale-invariant PPP.

Consider now the point process $\{(H_k, T_k), k = 1, 2, \ldots\}$ induced by the general stick-breaking (2). Applying (9, Theorem 1) yields:

Proposition 7. *If $m = \mathbb{E}[-\log W] < \infty$ then the scaled point process $\{(bH_k, T_k/b), k = 1, 2, \ldots\}$ has a weak limit $\mathcal{R}$ as $b \to \infty$. The limiting point process $\mathcal{R}$ in $\mathbb{R}_+^2$ is invariant under hyperbolic shifts $(s,t) \mapsto (bs, t/b)$ (with $b > 0$), and the coordinate projections of $\mathcal{R}$ are scale-invariant point processes.*

A more explicit construction of $\mathcal{R}$ is the following. Let $\mathcal{H}$ be the unique scale-invariant (that is, satisfying $b\mathcal{H} =_d \mathcal{H},\ b > 0$) point process on $\mathbb{R}_+$, whose restriction to $[0,1]$ has the same distribution as $\{W_0, W_0W_1, W_0W_1W_2, \ldots\}$ where the W_j's are independent, $W_k =_d W$ and W_0 has the stationary density (9). Let $\{\xi_k,\ k \in \mathbb{Z}\}$ be the points of $\mathcal{H}$ which may be labelled so that $\xi_0 = W_0$ is the maximum point of $\mathcal{H} \cap [0,1]$, and $\xi_{-1} > 1$. Assign to each ξ_k an arrival time $\sigma_k := \sum_{i=-\infty}^{k} E_i/\xi_i$ where the E_i's are independent standard exponential variables, also independent of $\mathcal{H}$. Then let $\mathcal{R} := \{(\xi_k, \sigma_k),\ k \in \mathbb{Z}\}$. The hyperbolic invariance of $\mathcal{R}$ is obvious from the construction and the scale invariance of $\mathcal{H}$.

Next is a counterpart of Proposition 5.

Proposition 8. *If one of the cooordinate projections of $\mathcal{R}$ is a Poisson process, then the law of W is* beta$(\theta, 1)$ *for some $\theta > 0$, in which case both projections are Poisson processes with intensity $\theta \mathrm{d}s/s,\ s > 0$.*

Proof. If the vertical coordinate projection of $\mathcal{R}$ is Poisson then the intensity is $\theta \mathrm{d}s/s$ for some $\theta > 0$, in consequence of scale invariance, but then W is beta$(\theta, 1)$ and the horizontal projection is the same Poisson, as is known in connection with the Ewens partition. The second part of the claim follows by observing that the law of the horizontal projection uniquely determines the law of W. Indeed, let $\xi_1 < \xi_2$ be two leftmost points of the horizontal projection on $[1, \infty[$. By the construction of $\mathcal{R}$ we have $\xi_1/\xi_2 =_d (E_1/E_2)W$, with exponential E_1, E_2 an E_1, E_2, W independent. Hence the law of W can be recovered from that of $\mathcal{R}$. □

In application to the chain records Proposition 7 guarantees point-process convergence of the heights and record times, but tells a little about the vector values of records. Observe that for the values of chain records we have the component-wise representation

$$R_k^{(j)} =_d U_1^{(j)} \cdots U_k^{(j)}, \quad j = 1, \ldots, d; \quad k = 1, 2, \ldots$$

with independent uniform $U_k^{(j)}$'s. From this it is clear that each marginal process $\{bR_k^{(j)}, k = 1, 2, \ldots\}$ converges to the same scale-invariant PPP on $\mathbb{R}_+$. However, the vector point process $\{bR_k\}$ converges, as $b \to \infty$, to a degenerate limit in $\mathbb{R}_+^d$ which lives on the union of the coordinate axis (this follows because any level c/b is surpassed by one of the marginal $\{R_k^{(j)}\}$'s considerably before the others).

As noted by Charles Goldie, the component-wise logarithmic transform

$$(-\log(R_k^{(1)}), \ldots, -\log(R_k^{(d)})), \quad k = 1, 2, \ldots$$

sends the chain records in Q_d to the sequence of sites visited by a d-dimensional random walk whose components are independent one-dimensional random walks with exponentially distributed increments. Equivalently, one can consider the *upper* chain records from the product exponential distribution in d dimensions. In this regime, R_k/k converge almost surely to the vector with unit coordinates and, subject to a suitable normalization, the process of record values can be approximated by a d-dimensional Brownian motion.

6 Chain records from general random orders

Let $(\mathcal{X}, \mathcal{F}, \mu, \prec)$ be a probability space endowed with a measurable strict partial order $\prec$. The measurability condition means that the graph of $\prec$ belongs to the product sigma-algebra $\mathcal{F} \otimes \mathcal{F}$, and it implies that the lower sections $L_x = \{y \in \mathcal{X} : y \prec x\}$ are all measurable and that the height $h(x) = \mu(L_x)$ is a measurable function from $\mathcal{X}$ to $[0,1]$. To avoid measure-theoretical pathologies like nonmeasurability of the diagonal we assume in the sequel that the space is sufficiently regular; one can think e.g. of a Borel subset of some Euclidean space. In this general setting, the chain records in a sample $X_1, X_2, \ldots$ from $(\mathcal{X}, \mathcal{F}, \mu, \prec)$ and the associated variables H_k, R_k, T_k, I_n are defined in full analogy with that in Q_d. We allow measures μ with atoms, hence some marks in the sample may be repeated.

Each measurable subset $A \subset \mathcal{X}$ of positive measure inherits the structure of a partially ordered probability space, endowed with the conditional measure $\mu_A := \mu(\cdot)/\mu(A)$ and the height function $h_A(x) := \mu(L_x \cap A)/\mu(A),\ x \in A$. If A is a *lower set*, meaning that with each $x \in A$ also $L_x \subset A$, then $h_A(x) = h(x)/\mu(A)$. Note that by transitivity each L_x is a lower set, and every union of lower sets is a lower set.

The next lemma states that the height function restricted to the sample is strictly monotone.

Lemma 9. *We have* $\mathbb{P}((X_1 \prec X_2) \wedge (h(X_1) \not< h(X_2)) = 0$.

Proof. By transitivity, $x \prec y$ implies $L_x \subset L_y$ and $h(x) \leq h(y)$ for almost all (x, y). Thus we only need to show that the event $X_1 \prec X_2$ excludes $h(X_1) = h(X_2)$. Suppose this does not hold, then $\mathbb{P}((X_1 \prec X_2) \wedge (h(X_1) = h(X_2))) > 0$. Then by Fubini there is a set A of positive measure such that for $x \in A$ we have $\mathbb{P}((X_1 \prec x) \wedge (h(X_1) = h(x))) > 0$, and since in this event $L_{X_1} \subset L_x$ we have $\mathbb{P}((X_1 \prec x) \wedge (L_{X_1} = L_x)) > 0$. By exchangeability the latter is also true with X_2 substituted for X_1, hence by independence $\mathbb{P}((X_1 \prec x,\ X_2 \prec x) \wedge (L_{X_1} = L_{X_2} = L_x)) > 0$ for almost all $x \in A$. This implies

$$\mathbb{P}((X_1 \in L_{X_2}) \wedge (X_2 \in L_{X_1})) = \mathbb{P}((X_1 \prec X_2) \wedge (X_2 \prec X_1)) > 0,$$

which contradicts to the asymmetry of $\prec$. □

It follows that (H_k) is strictly decreasing almost surely. Lemma 2 is still valid in the general framework, but the law of (H_k) typically depends in a very complex way on the sampling space. As the first instance, we discuss features of H_1 and relate them to the properties of the chain records. Introduce the distribution function

$$D(s) := \mathbb{P}(H_1 \leq s) = \mathbb{P}(h(X_1) \leq s) = \mu\{x \in \mathcal{X} : h(x) \leq s\}, \quad s \in [0,1].$$

In a standard way, D is right-continuous with left limits, nondecreasing and satisfies $D(1) = 1$ and $D(0) \geq 0$.

Proposition 10. *The total number of chain records is almost surely infinite if and only if* $D(0) = 0$.

Proof. If $D(0) = 0$ the process (X_n) never enters the set $\{x \in \mathcal{X} :\ h(x) = 0\}$ hence each chain record is eventually broken, and the sequence of chain records is infinite. If $D(0) > 0$ then with positive probability $h(X_1) = 0$ and the record process terminates at $T_1 = 1$. □

Example 11. The distribution D may have jumps even when μ has no atoms. An example of this situation is $\mathbb{R}^2$ with the standard order and with the measure supported by the union of the unit square and the segment connecting points $(1, 0)$ and $(2, -1)$, such that a half of the mass is spread uniformly over the square and another half is spread uniformly over the segment. Observe that $D(0) = 1/2$, with probability $1/2$ the total number of chain records is infinite, and with probability $1/2$ it is finite, depending on whether X_1 hits the square or the segment.

Lemma 12. *The distribution function D is superdiagonal: $D(s) \geq s$ for $s \in [0, 1]$.*

Proof. We show first the continuity at 1. Suppose $D(1) - D(1-) = \alpha > 0$ and let $C = \{x \in \mathcal{X} : h(x) = 1\}$. By Fubini and exchangeability

$$\mathbb{P}((X_2 \not\prec X_1) \wedge (X_1 \in C)) = \mathbb{P}((X_1 \not\prec X_2) \wedge (X_2 \in C)) = 0.$$

This readily implies

$$\mathbb{P}((X_1 \prec X_2) \wedge (X_2 \prec X_1)) \geq \mathbb{P}((X_1 \prec X_2) \wedge (X_2 \prec X_1) \wedge (X_1 \in C) \wedge (X_2 \in C)) = \\ \mathbb{P}((X_1 \in C) \wedge (X_2 \in C)) = \alpha^2 > 0.$$

This contradicts to the asymmetry of partial order, thus $D(1) = D(1-)$.

Suppose now that s is a growth point for D. Then for every ϵ there is a set $A \subset \mathcal{X}$ of positive measure such that $h(x) \in [s - \epsilon, s + \epsilon]$ for $x \in A$. Hence

$$D(s + \epsilon) = \mathbb{P}(h(X_2) \leq s + \epsilon) \geq \mathbb{P}(h(X_2) \leq h(X_1) \,|\, X_1 \in A) \geq \mathbb{P}(X_2 \prec X_1 \,|\, X_1 \in A) \geq s - \epsilon\,,$$

and letting $\epsilon \to 0$ we obtain $D(s) \geq s$.

It remains to exclude the possibility that D has a flat which ends with crossing the diagonal by a jump. Suppose $s - D(s-) > 0$ for some s. Then $B = \{x \in \mathcal{X} : h(x) \leq s\}$ is a lower set of positive measure, whose height function h_B has distribution with an atom at 1. This is impossible by the first part of the argument. $\square$

Example 13. (minimal orders) For arbitrary superdiagonal D there exists an ordered probability space such that the distribution function of $h(X_1)$ is D. Indeed, as noted in (6, p. 85) for the partial order on $([0, 1], \mathcal{B}, \mathrm{d}x)$ defined by

$$x \prec y \iff D(x) < y\,, \quad \text{for } x, y \in [0, 1]. \tag{12}$$

the distribution function of $h(X_1)$ is D, because $L_x = [0, D^{\leftarrow}(x)[$. The relation defined by (12) may be called a *minimal order*, since it is a minimal (in the sense of set-theoretic inclusion modulo μ) partial order on $([0, 1], \mathcal{B}, \mathrm{d}x)$ with given distribution D for the height function. For the minimal orders, the strong records coincide with the chain records, and the number of chain records is finite or infinite depending on whether $D(0) > 0$ or $D(0) = 0$.

Example 14. The space of reals $(\mathbb{R}, \mathcal{B}, \mu, <)$ with arbitrary μ is isomorphic to a minimal space in the previous example. The isomorphism is achieved by the probability transform $x \mapsto \mu] - \infty, x]$. If μ has no atoms then $D(s) = s$ is the uniform distribution, and in general D is a *uniformized* distribution (22), with the characteristic property that $D(s-) = s$ for D-almost all $s \in [0, 1]$.

Partially ordered probability spaces $(\mathcal{X}, \mathcal{F}, \prec, \mu)$ for which D is uniformized have the characteristic property that the event $\{X_1 \prec X_2\}$ coincides with $\{h(X_1) < h(X_2)\}$. In this sense the height h plays the role of a 'utility function' representing the relation $\prec$ restricted to the sample. Another characterization of this situation is that the relation $\approx$ defined by $x \approx y \Longleftrightarrow (x \not\prec y) \wedge (y \not\prec x)$ is transitive modulo μ.

There is a canonical minimal order $\prec$ on $([0,1], \mathcal{B}, \mathrm{d}x)$ (see previous example) related to $(\mathbb{R}, \mathcal{B}, \mu, <)$. For $x, y \in [0,1]$ let $x \prec y$ iff $D(x) < D(y)$. This makes two distinct point x and y incomparable if they belong to the same flat of D. Each flat of D corresponds to some atom of μ, and has the length equal to the size of this atom.

7 Comparing the records of distinct types

In the setting of Section 6 one can define also weak and strong records, as in the Introduction for Q_d. Let $\underline{p}_n, \overline{p}_n, p_n$ be the probabilities that a weak, a strong or a chain record occurs at time n, and let $\underline{N}_n, \overline{N}_n, N_n$ be the number of records among the first n marks for each of these types, respectively. Obviously,

$$\mathbb{E}\,[\underline{N}_n] = \sum_{j=1}^{n} \underline{p}_n, \quad \mathbb{E}\,[\overline{N}_n] = \sum_{j=1}^{n} \overline{p}_n, \quad \mathbb{E}\,[N_n] = \sum_{j=1}^{n} p_n\,.$$

The expected total number of records of each kind in the infinite sequence $X_1, X_2, \ldots$ is given by the corresponding infinite series. For weak and strong records, the divergence of the series is a necessary and sufficient condition that infinitely many records occur almost surely, see (30, Theorem 2.1). The probabilities $\underline{p}_n, \overline{p}_n$ and p_n are nonincreasing with n, as it follows by observing that a record occurs when X_n hits a certain subset of $\mathcal{X}$ determined by $X_1, \ldots, X_{n-1}$, and these subsets are nonincreasing.

By definitions and exchangeability $\underline{N}_n \le N_n \le \overline{N}_n$,

$$\underline{p}_n \le p_n \le \overline{p}_n \quad \text{and} \quad \underline{p}_n \le 1/n \le \overline{p}_n\,. \tag{13}$$

Complementing these relations we will show the inequality

$$\mathbb{E}\,[N_n] \le \sum_{j=1}^{n} 1/j\,, \tag{14}$$

which makes a sense of the statement that the chain records are more rare than the standard records from Q_1. As a nice exercise, the reader is suggested to prove (14) for strict lower records in a sample from some discrete distribution on $\mathbb{R}$. For the probability of a weak record we have the integral representation

$$\overline{p}_n = \int_0^1 (1-s)^{n-1}\,\mathrm{d}D(s)\,, \tag{15}$$

which follows by conditioning on $h(X_n) = s$. A similar representation exists for $\underline{p}_n$ in terms of the distribution of the upper section of $\prec$ at a random mark X_1, but there is no general simple formula for p_n.

Introduce the ratios

$$W_k = H_k/H_{k-1}\,, \quad k = 1, 2, \dots$$

with the convention that $H_0 = 1$ and $H_k = W_k = 0$ if the number of chain records is less than k. Thus $H_k = W_1 \cdots W_k$. Let (U_k) be a sequence of independent uniform $[0, 1]$ random variables.

Lemma 15. *It is possible to define (H_k) and (U_k) on the same probability space to achieve $W_k \le U_k$ hence $H_k \le U_1 \cdots U_k$ for all $k = 1, 2, \dots$.*

Proof. Let D_x be the distribution of height for the lower set L_x with conditional measure $\mu(\cdot)/\mu(L_x)$. If $\mu(L_x) = 0$ we let $D_x \equiv 1$. Let $D^{\leftarrow}$ and $D_x^{\leftarrow}$ be the generalized inverses of D and D_x, respectively. By Lemma 12 we have $D_x(s) \ge s$ thus $D_x^{\leftarrow}(u) \le u,\ u \in [0, 1]$. Let (X_n) and (U_k) be independent. Define $W_1' := D^{\leftarrow}(U_1)$ and $W_{k+1}' := D_{R_k}^{\leftarrow}(U_{k+1})$ for $k \ge 1$. By properties of the quantile transform we have the distributional identity

$$W'_{k+1} \,|\, R_k {=}_d W_{k+1} \,|\, R_k\,,$$

and the same is true for conditioning on $R_1, \dots, R_k$ since (R_k) is Markovian. It follows by induction in k that $(W_k'){=}_d(W_k)$. Because $W_k' \ge U_k$ we are done. □

Proposition 16. *For N_n^* the number of classical records from Q_1 and N_n the number of chain records from a general space $\mathcal{X}$ we have*

$$\mathbb{P}(N_n \ge k) \le \mathbb{P}(N_n^* \ge k).$$

In particular, (14) *is true.*

Proof. Let $H_k^* = U_1 \cdots U_k$, $H_0^* = 1$. Suppose that (T_k^*) is such that given (H_k^*) all $T_{k+1}^* - T_k^*$ are independent geometric(H_k^*), and $T_1^* = 1$. Define $N_n^* := \max\{k : T_k^* \le n\}$. Then (H_k^*), (T_k^*) and N_n^* may be identified with the record-values, records times and record counts for the standard records from Q_1. In the setup of Lemma 15 we have $H_k \le H_k^*$, hence T_k^* is stochastically smaller than T_k, hence each N_n^* is stochastically larger than N_n. □

8 Self-similar spaces

The special feature which makes the occurences of chain records from Q_d lending itself to study is the stick-breaking representation (2) with i.i.d. factors, which in turn follows from a self-similarity property of the cube. This property holds for a wider class of spaces introduced below.

We define a partially ordered probability space $(\mathcal{X}, \mathcal{F}, \mu, \prec)$ to be *self-similar* if for almost all $x \in \mathcal{X}$ the conditioned space $(L_x, \mathcal{F}|_{L_x}, \mu_{L_x}, \prec|_{L_x})$ is isomorphic to the whole space $(\mathcal{X}, \mathcal{F}, \mu, \prec)$. The isomorphism is understood as a measure-preserving bijection ϕ defined almost everywhere and such that the events $\{X_1 \prec X_2\}$ and $\{\phi(X_1) \prec \phi(X_2)\}$ coincide up to a zero event. For a self-similar space the factors in (2) are i.i.d. copies of W whose distribution is D.

Bollobás and Brightwell (12) introduced *box-spaces* with the property that all intervals $\langle y, x\rangle := \{z \in \mathcal{X} : y \prec z \prec x\}$ $(y \prec x)$ are isomorphic to the whole space. Let o be the unique least element of $\mathcal{X}$ (if $\mathcal{X}$ has no such element we attach o to the space), then $L_x = \langle o, x\rangle$. Thus our definition of self-similar space requires isomorphism of a smaller class of intervals, hence every box-space is self-similar.

We demonstrate next various constructions of self-similar spaces.

Example 17. ($\mathbb{R}^d$ as a product space) The cardinal example is, of course, $\mathbb{R}^d$ with coordinate-wise order and a continuous product distribution $\mu = \otimes_{j=1}^d \mu^{(j)}$. The transformation $(x^{(1)}, \dots, x^{(d)}) \mapsto (\mu^{(1)}]-\infty, x^{(1)}], \cdots, \mu^{(d)}]-\infty, x^{(d)}])$ is an isomorphism with Q_d, thus $\mathbb{R}^d$ with such measure is indeed a self-similar space.

The family of self-similar spaces is closed under the operation of cartesian product, with the product order defined analogously to the case $Q_d = [0,1] \times \cdots \times [0,1]$.

Example 18. (discrete measures) For (z_k) a decreasing sequence and $0 < p < 1$, $q = 1-p$ let $\mu = \sum_{k=1}^\infty pq^{k-1}\delta_{z_k}$, where δ_x is the Dirac mass at x. Clearly, $h(z_k) = q^k$. The probability transform yields the measure of the same type, but with supporting sequence $z'_k = q^k$. The space $(\mathbb{R}, <)$ with any such measure is self-similar, as it follows from the elementary properties of the geometric distribution. Computing the moments m and σ^2 is easy from the Mellin transform $g(\lambda) = pq^\lambda/(1-q^{\lambda+1})$. From the moments formula

$$\mathbb{E}[Y^k] = \frac{k!}{m}\frac{1-q^k}{k}, \qquad m = -\log q$$

follows that Y is a product of independent factors: exponential(1) and a variable with density $\mathrm{d}s/(ms)$ $s \in [q,1]$. For $q \to 1$ the second factor degenerates and the distribution of Y converges to exponential(1), as is intuitively clear since μ approaches the uniform distribution on $[0,1]$, hence the records must be similar to that from Q_1.

If $(\mathbb{R}, \mathcal{B}, \mu, <)$ is self-similar then μ is either continuous or it is geometric as in Example 18. Indeed, applying the probability transform we reduce to the case $\mu[0,s] = D(s)$ of a uniformized measure μ. From the fact that H_2/H_1 and H_1 are i.i.d. and (27, Corollary 6.5) follows that the support of μ is a multiplicatively regenerative set, but any such nonrandom set is either $[0,1]$ or a geometric sequence, as is obvious from (27, Section 6).

Bunge and Goldie (8, p. 297) asked if the Ewens sampling formula with $\theta \neq 1$ may be given some interpretation in terms of records. In (28) an interpretation was given for arbitrary $\theta > 0$ in terms of the classical records and a 'F^α-model' for independent (but not i.i.d.) marks. The next example gives an interpretation in terms of suitable chain records, but only for $0 < \theta \leq 1$. By Lemma 12 the parameter values $\theta > 1$ cannot appear in this way, since then beta$(\theta, 1)$ is subdiagonal and cannot coincide with some height distribution.

Example 19. (unit square with a weird order) Consider the square $[0,1]^2$ with μ being the product of the uniform distribution $[0,1]$ and a beta$(\beta,1)$ distribution ($\beta > 0$). Thus $\mu([0,x^{(1)}] \times [0,x^{(2)}]) = x^{(1)}(x^{(2)})^\beta$ for $x = (x^{(1)}, x^{(2)}) \in [0,1]^2$. Define the partial order

$$(x^{(1)}, x^{(2)}) \prec (y^{(1)}, y^{(2)}) \iff \max(x^{(1)}, x^{(2)}) < y^{(1)}.$$

Thus each L_x for $x = (x^{(1)}, x^{(2)}) \in [0,1]$ is itself a square $[0, x^{(1)}[\times[0, x^{(1)}[$, and appealing to the elementary properties of the beta distributions we readily derive self-similarity of the ordered sampling space. Noting that $h(x) = (x^{(1)})^{\beta+1}$ it follows that $D(s) = s^{1/(1+\beta)}$ is the beta$(1/(1+\beta), 1)$ distribution. Proposition 5 applies with $\theta = 1/(\beta+1)$, and the indicators of chain records are independent. If we view the sequence of n indicators as a composition of n (e.g. 1010000100 corresponds to $(2,5,3)$), then this composition follows the Ewens sampling formula with parameter θ. The scaled point process of record times $\{T_k/n,\ k=1,2,\ldots\}$ converges to a scale-invariant Poisson process. We have $g(\lambda)=\theta/(\theta+\lambda)$ and moment computations show that Y is gamma(θ).

Analogous examples with the independence property of indicators can be also constructed in higher dimensions.

Example 20. (interval spaces) A random interval $]a,b[\subset [0,1]$ is defined by choosing its endpoints $0\le a<b\le 1$ from the bivariate density $\alpha(\alpha-1)(b-a)^{\alpha-2}$, where $\alpha>1$. For $\alpha=2$ this is a bivariate uniform distribution. One natural partial order on the intervals is by inclusion. Then

$$D'(s)=(\alpha-1)(s^{-1/\alpha}-1)\,,\quad g(\lambda)=\frac{\alpha-1}{(\lambda+1)(\alpha\lambda+\alpha-1)}. \tag{16}$$

The limit variable is $Y=_d EZ$ with independent E exponential(1) and Z beta$(1-1/\alpha,1)$.

Another partial order is the 'interval order' defined by $]a,b[\prec]a',b'[\iff b\le a'$. With this order and the above density the space of intervals is a box-space (12, p. 63). For this space we have

$$D'(s)=\left[s^{-1/\alpha}-1\right]^{\alpha-1}\,,\quad g(\lambda)=\frac{\Gamma(1+\alpha)\Gamma(1+\alpha\lambda)}{\Gamma(1+\alpha\lambda+\lambda)}.$$

Both constructions of interval spaces generalize to arbitrary box-spaces.

Example 21. (cones) Let $H\subset\mathbb{R}^d$ be a hyperplane at unit distance from the origin and let A be a $(d-1)$-dimensional convex body in this hyperplane. Define a self-similar space with $\mathcal{X}$ being the convex hull of $A\cup\{0\}$, with μ being the measure with density $c\rho(x)^\beta$ ($\beta>-1$) where $\rho(x)$ is the distance from x to H, and with order $x\prec y\iff x-y\in\mathcal{X}$ for $x,y\in\mathcal{X}$. It is easy to compute $h(x)=\rho(x)^{\beta+d}$ and

$$D'(s)=c\,s^{(1-d)/(\beta+d)}(1-s^{1/(\beta+d)})^{d-1},$$

where $c=(\beta+1)\cdots(\beta+d-1)/(d-1)!$.

The last formula specializes to (16) for $\beta=\alpha-2$, $d=2$. This is no coincidence: an inspection of the interval space shows that it fits in the present example.

Example 22. (simplexes) This example is suggested by a construction of simplex trees in (14). Let $\mathcal{X}$ be a d-dimensional simplex with μ the uniform distribution on $\mathcal{X}$. Fix a $(d-1)$-dimensional face of $\mathcal{X}$ and for $x\in\mathcal{X}$ let L_x be the interior of the convex hull of $x\cup F$. Define a partial order by $y\prec x\iff y\in L_x$. The law of W is $D(s)=1-(1-s)^d$. This distribution is beta$(1,d)$ ($d=1,2,\ldots$), but the parameter stands here on the inappropriate side, hence the record indicators are independent only for $d=1$.

Bollobás and Brightwell (12) also defined a larger class of spaces which have all intervals isomorphic to one another. By analogy, we can introduce a larger class of spaces which have (almost) all lower sections isomorphic to one another, but not necessarily isomorphic to the whole space. Examples are easy to provide, one can take for instance any lower set in a self-similar space. For this larger class of spaces (2) still holds with independent factors, but only $W_2, W_3, \ldots$ must be i.i.d. The CLT for the number of chain records readily extends to this more general situation.

Acknowledgment The author is indebted to Charles Goldie for the remark in Section 5, comments on an earlier draft of the paper and his most stimulating interest.

References

[1] Arnold, B.C., Balakrishnan, N. and Nagaraja, H.N. (1998) *Records*, Wiley, New York. MR1628157

[2] Arratia, R., Barbour, R.D. and Tavaré, S. (1999) The Poisson-Dirichlet distribution and the scale invariant Poisson process, *Combin. Prob. Comp.* **8** 407-416. MR1731976

[3] Bai, Z.-D., Chao, C.-C., Hwang, H.-K. and Liang, W.-Q. (1998) On the variance of the number of maximal vectors and its applications, *J. Appl. Probab.* **8**, 386-395 MR1627803

[4] Barndorff-Nielsen, O. and Sobel, M. (1966) On the distribution of the number of admissible points in a vector random sample, *Th. Probab. Appl.* **11**, 283-305. MR0207003

[5] Baryshnikov, Y.M. (2000) Supporting points processes and some of their applications, *Prob. Th. Rel. Fields* **117**, 163-182. MR1771659

[6] Berezovsky, B.A., Baryshnikov, Y. M., Borzenko, V.I. and Kempner, L.M. (1989) *Multicriteria optimization : mathematical aspects*, Moscow, Nauka (in Russian). MR1019318

[7] Bertoin, J. (2006) *Random fragmentation and coagulation processes*, Cambridge University Press. MR2253162

[8] Bunge, J. and Goldie, C.M. (2001) Record sequences and their applications, In *Handbook of Statistics* vol. 19 (Shanbhag, D.N. and Rao, C.R. eds) 277-308. MR1861727

[9] Bertoin, J. and Caballero, M.-E. (2002) Entrance from 0+ for increasing semi-stable Markov processes, *Bernoulli* **8**, 195-205. MR1895890

[10] Bertoin, J. and Gnedin, A.V. (2004) Asymptotic laws for nonconservative self-similar fragmentations, *Elect. J. Probab.* **9**, 575-593. MR2080610

[11] Bertoin, J. and Yor, M. (2005) Exponential functionals of Lévy processes, *Probability Surveys* **2**, 191-212. MR2178044

[12] Bollobas, B. and Brightwell, G. (1991) Box-spaces and random partial orders, *Trans. Amer. Math. Soc.* **324**, 59-72. MR0986685

[13] Deuschel, J.-D. and Zeitouni, O. (1995) Limiting curves for iid records, *Ann. Probab.* **23**, 852-878. MR1334175

[14] Devroye, L. (1999) Universal limit laws for depths in random trees, *SIAM J. Comp.* **28**, 409-432. MR1634354

[15] Devroye, L. and Laforest, L. (1990) An analysis of random d-dimensional quad trees, *SIAM J. Comput.* **19**, 821-832. MR1059656

[16] Feller, W. (1971) *An Introduction to probability theory and its applications*, vol. 2, Wiley, NY. MR0270403

[17] Flajolet, P., Gonnet, G., Puech, C. and Robson, J.M. (1993) Analytic variations on quadtrees, *Algorithmica* **10**, 473-500. MR1244619

[18] Flajolet, P., Labelle, G., Laforest, L. and Salvy, B. (1995) Hypergeometrics and the cost structure of quadtress, *Random Struct. Alg.* **7**, 117-144. MR1369059

[19] Gnedin, A. (1994) On a best-choice problem with dependent criteria, *J. Appl. Prob.* **31** 221-234. MR1260584

[20] Gnedin, A. (1994) Conical extremes of a multivariate sample *J. Res. Natl. Inst. Stand. Techn.* **99**, 511-519.

[21] Gnedin, A. (1993) On multivariate extremal processes, *J. Multivar. Anal.* **46**, 207-213. MR1240421

[22] Gnedin, A. (1997) The representation of composition structures, *Ann. Probab.* **25**, 1437-1450. MR1457625

[23] Gnedin, A. (2004) Best choice from the planar Poisson process, *Stoch. Proc. Appl.* **111**, 317-354. MR2056541

[24] Gnedin, A. (2007) Recognising the last record of a sequence, *Stochastics* **79**, 199-209.

[25] Gnedin, A. (2004) The Bernoulli sieve, *Bernoulli* **10**, 79-96. MR2044594

[26] Gnedin, A. (1998) Records from a multivariate normal sample, *Statist. Prob. Lett.* **39**, 11-15. MR1649331

[27] Gnedin, A. and Pitman, J. (2005) Regenerative composition structures, *Ann. Probab.* **33**, 445-479. MR2122798

[28] Gnedin, A. and Pitman, J. (2005) Markov and selfsimilar composition structures, *Zapiski POMI (St. Petersburg Dept. Steklov Math. Inst.* **326**, 59-84. MR2183216

[29] Gnedin, A. and Pitman, J. (2007) Poisson representation of a Ewens fragmentation process, *Comb. Prob. Comp.*

[30] Goldie, C.M. and Resnick, S. (1989) Records in a partially ordered set, *Ann. Probab.* **17**, 678-699. MR0985384

[31] Goldie, C.M. and Resnick, S. (1995) Many multivariate records, *Stoch. Process. Appl.* **59**, 185-216. MR1357651

[32] Goldie, C. M. and Resnick, S. (1996) Ordered independent scattering, *Comm. Stat., Stoch. Models* **12**, 523-528. MR1410845

[33] Golin, M. (1993) Maxima in convex regions, *Proc. Fourth Annual ACM-SIAM Symp. on Discrete Algorithms (Austin, TX, 1993)*, ACM, New York, 352-360 MR1213247

[34] Hashorva, E. and Hüsler, J. (2002) On asymptotics of multivariate integrals with applications to records, *Stoch. Models* **18**, 41-69. MR1888285

[35] Hashorva, E. and Hüsler, J. (2005) Multiple maxima in multivariate samples, *Stoch. Prob. Letters* **75**, 11-17. MR2185605

[36] Naću, S. (2006) Increments of random partitions, *Comb. Prob. Comp.* **15**, 589-595. MR2238047

[37] Nevzorov, V.B. (2001) *Records: mathematical theory*, Transl. Math. Monographs **194**, Amer. Math. Soc., Providence RI. MR1791071

[38] Pitman, J. (1995) Exchangeable and partially exchangeable random partitions, *Prob. Th. Rel. Fields* **102**, 145-158. MR1337249

[39] Pitman, J. (2006) *Combinatorial stochastic processes*, Springer L. Notes Math. **1875**. MR2245368