The Annals of Probability
1989, Vol. 17, No. 2, 678–699

RECORDS IN A PARTIALLY ORDERED SET

BY CHARLES M. GOLDIE AND SIDNEY RESNICK[1]

University of Sussex and Cornell University

We consider independent identically distributed observations taking values in a general partially ordered set. Under no more than a necessary measurability condition we develop a theory of record values analogous to parts of the well-known theory of real records, and discuss its application to many partially ordered topological spaces. In the particular case of $\mathbb{R}^2$ under a componentwise partial order, assuming the underlying distribution of the observations to be in the domain of attraction of an extremal law, we give a criterion for there to be infinitely many records.

1. Introduction. Let S be a set equipped with a strict partial order $<$, that is, $\{(x, y): x < y\}$ is a subset of $S \times S$ such that the properties of

antisymmetry: $x < x$ for no $x \in S$,
transitivity: if $x < y$ and $y < z$, then $x < z$

hold in S. We shall set up a sequence of i.i.d. (independent identically distributed) S-valued random elements $X_1, X_2, \dots$ and say that

(1.1) $$X_n \text{ is a } \textit{record} \text{ iff } X_k < X_n, \qquad k = 1, \dots, n-1.$$

To motivate and inform what follows, consider the case of i.i.d. random vectors $X_n = (X_n^{(1)}, X_n^{(2)})$ in $\mathbb{R}^2$. The following notions are all plausible:

(i) X_n is a "record" iff $X_n^{(1)} > \bigvee_{k=1}^{n-1} X_k^{(1)}$ and $X_n^{(2)} > \bigvee_{k=1}^{n-1} X_k^{(2)}$;
(ii) X_n is a "record" iff $X_n^{(1)} \geq \bigvee_{k=1}^{n-1} X_k^{(1)}$ and $X_n^{(2)} \geq \bigvee_{k=1}^{n-1} X_k^{(2)}$ and at least one of these inequalities is strict;
(iii) X_n is a "record" iff $X_n^{(1)} > \bigvee_{k=1}^{n-1} X_k^{(1)}$ or $X_n^{(2)} > \bigvee_{k=1}^{n-1} X_k^{(2)}$;
(iv) X_n is a "record" iff X_n falls outside the convex hull of $X_1, \dots, X_{n-1}$;
(v) X_n is a "record" iff $|X_n| > \bigvee_{k=1}^{n-1} |X_k|$,

where $|\cdot|$ in (v) denotes Euclidean (or supremum, or some other) norm. Definition (iii) may be the most natural from the statistical point of view while (iv) perhaps has the most intrinsic mathematical depth. However, these two are not reconcilable with (1.1) for any partial order on $\mathbb{R}^2$, and we shall not deal with them here. We shall, though, cover (i) and (ii) as part of our general framework, and obtain some nice probabilistic structure which is absent from (iii) and (iv). A fortiori, our results will apply to (v), but there much more is known because (v)

Received November 1986; revised May 1988.
[1]Supported during 1985–1986 by a UK Science and Engineering Research Council grant. The hospitality of the Mathematics Division, University of Sussex, is acknowledged. Additional partial support came from NSF Grant DMS-85-01763.
AMS 1980 *subject classifications.* Primary 60B05; secondary 60K99, 06A10.
Key words and phrases. Bivariate extremal law, continuous lattice, Fell topology, hazard measure, lattice, Lawson topology, partially ordered set, random closed set, records, semicontinuity, sup vague topology, upper semicontinuity.

determines the "records" by the total order on the real-valued *attribute* $|\cdot|$, and most of the well-known theory of records of real i.i.d. observations carries over [Deken (1976) and Goldie and Rogers (1984)]. In what follows we find analogues of parts of the theory of real records in the setting of (1.1).

Some notation: It is standard [Birkhoff (1967), page 1] that $<$ determines a weak partial order $\leq$ by

$$x \leq y \quad \text{iff } x < y \text{ or } x = y,$$

that $\leq$ has the usual properties of reflexivity ($x \leq x\ \forall\ x$), antisymmetry ($x \leq y$, $y \leq x$ imply $x = y$) and transitivity ($x \leq y$, $y \leq z$ imply $x \leq z$) and that $\leq$ determines $<$, with correct properties, in the obvious way. Thus it is indifferent whether we start with $<$ or $\leq$.

We adjoin to S two points $-\infty, \infty$ with the properties

$$-\infty < x < \infty, \quad \text{all } x \in S,$$

and write $S^* := S \cup \{-\infty, \infty\}$. We need to do this whether or not S already has a greatest or least element. We use $+\infty$ as a "cemetery" state for the record-value sequence, in case the sequence of records in S terminates, and we use $-\infty$ as a fictitious starting state, before the records in S begin. Observe that $<$ and $\leq$ extend to S^*.

Define *intervals*

$$(x, y) := \{z \in S^*: x < z < y\}, \qquad [x, y) := \{x\} \cup (x, y),$$

and so on for $(x, y], [x, y]$.

Let $\mathscr{S}$ be a σ-algebra of subsets of S. We form $\mathscr{S}^*$, a σ-algebra of subsets of S^*, by adjoining to $\mathscr{S}$ the singletons $\{-\infty\}, \{\infty\}$ and their unions with elements of $\mathscr{S}$. Let μ be a probability measure on $(S, \mathscr{S})$. We extend it to $(S^*, \mathscr{S}^*)$ by setting $\mu\{-\infty\} := 0 =: \mu\{\infty\}$. Let $(\Omega, \mathscr{A}, P)$ be the product probability space $(\Pi S_n^*, \Pi \mathscr{S}_n^*, \Pi \mu_n)$, where the products run over $n \in \mathbb{N}$, and $(S_n^*, \mathscr{S}_n^*, \mu_n)$ are copies of $(S^*, \mathscr{S}^*, \mu)$. The generic element of Ω is $\omega = (\omega_1, \omega_2, \dots)$, and for each $n \in \mathbb{N}$ we let $X_n: \Omega \to S^*$ be the nth coordinate projection,

$$X_n(\omega) := \omega_n, \qquad \omega \in \Omega.$$

Recall that the product σ-algebra $\Pi \mathscr{S}_n^*$ is the smallest that makes the coordinate projections measurable. Thus the X_n are (canonical) i.i.d. S^*-valued random elements, each with law μ supported by S.

Define the *graph* of $<$ as the set

$$G_< := \{(x, y) \in S \times S: x < y\}.$$

We need just one assumption connecting the partial order to the measure structure, namely that $G_<$ is product-measurable:

$$\text{(A)} \qquad G_< \in \mathscr{S} \times \mathscr{S}.$$

Hence each set $\{\omega: X_m(\omega) < X_n(\omega)\}$, for $m \neq n$, is an event (element of $\mathscr{A}$). Clearly, we need assumption (A) to be able to define records at all.

The x-sections of $G_<$, $\{y \in S: x < y\}$, are the intervals $(x, \infty) \subset S$, and the y-sections the intervals $(-\infty, y) \subset S$. All these are in $\mathscr{S}$ [Halmos (1950), Section 34, Theorem A]. So all sets $\{X_n < x\}, \{X_n > x\}$, for $n \in \mathbb{N}$, $x \in S$, are events.

Next, Fubini's theorem [in the form, e.g., in Halmos (1950), Section 35, Theorem A] says that $\int_S 1_{G_<}(x, y)\, d\mu(x) = \mu(-\infty, y)$ is measurable as a function of y. We let

$$S_\mu := \{y \in S: \mu(-\infty, y) < 1\};$$

then $S_\mu \in \mathscr{S}$. This set will perform the role of the support of μ as a consequence of the following result, which was independently proved by G. O'Brien and E. Perkins, and by W. Vervaat and an anonymous referee.

PROPOSITION 1.1. *Assume* (A). *Then* $\mu(S_\mu) = 1$.

PROOF. Write $C := S_\mu^c$. For $x \in C$ the set $(-\infty, x)$ is μ-full so $\mu(C \cap (-\infty, x)) = \mu(C)$. Integrate:

$$\int_C \mu(C \cap (-\infty, x))\, d\mu(x) = \mu^2(C);$$

that is,

$$\int_C \int_C 1\{y < x\}\, d\mu(y)\, d\mu(x) = \mu^2(C).$$

By Fubini and change of notation,

$$\int_C \int_C 1\{x < y\}\, d\mu(y)\, d\mu(x) = \mu^2(C).$$

The sum of the last two left-hand sides is at most $\mu^2(C)$. So $\mu^2(C) = 0$. □

We thank W. Vervaat also for the following remark: (A) is more general than measurability of $G_\le$. For when $G_\le$ is measurable so is $G_< = G_\le \setminus (G_\le \cap G_\ge)$. But for instance when S is a nonseparable metric space with its Borel σ-algebra, $G_< := \varnothing$ is measurable while $G_\le$ is not.

As we assume no more than (A) our results are of unrestricted generality. No lattice assumptions are needed on the partial order, no topology on S.

In particular cases, (A) needs to be checked, and the natural way to establish the link between partial order and measure that it represents is through an appropriate topology on S. We shall prove (A) for many partially ordered topological spaces in Section 4. As a consequence, it will hold in particular for the following spaces.

(a) $\mathbb{R}^d$ with $x = (x^{(1)}, \dots, x^{(d)}) < y = (y^{(1)}, \dots, y^{(d)})$ iff $x^{(j)} < y^{(j)}$ for $j = 1, \dots, d$.

(b) $\mathbb{R}^d$ with $x \le y$ iff $x^{(j)} \le y^{(j)}$ for $j = 1, \dots, d$. Here, $x < y$ iff $x^{(j)} \le y^{(j)}$ for $j = 1, \dots, d$, with strict inequality for some j.

(c), (d) $\mathbb{R}^{\mathbb{N}}$ ordered as in (a), respectively (b), above.

(e) $C[0,1]$ with $x \le y$ iff $x(t) \le y(t)$ for all $t \in [0,1]$.

(f) $\mathscr{F}(E)$, the space of closed sets in E, a locally quasicompact space with countable base (*lqccb*). The topology on $\mathscr{F}(E)$ is the Fell topology (see Section 4) and the weak partial order is set-inclusion. (By "quasicompact" we mean that every open cover has a finite subcover.)

(g) $US(E)$, the space of usc (upper semicontinuous) $[-\infty, \infty]$-valued functions on E, which is to be some *lqccb* set. The topology on $US(E)$ is the sup vague topology of Vervaat (1966, 1988b) and the weak partial order is as in (e) above.

(h) $\mathscr{P}$, the space of probability measures on $\mathbb{R}$. The "narrow" topology is appropriate, and the partial order can be one of a number of stochastic orderings of interest (see Section 4).

Observe that in (e) and (g), the random function X_n is a record iff $X_n(t) \ge \bigvee_{k=1}^{n-1} X_k(t)$ for all $t \in E$, with strict inequality for some t. And in (f) the random set X_n is a record iff $\bigcup_{k=1}^{n-1} X_k$ is a proper subset of X_n.

Our results about records, assuming (A), are in Sections 2 and 3 and may be summarized as follows. In Section 2 we define a measure H on $(S, \mathscr{S})$ called a *hazard measure*. $H(A)$ turns out to be the expected number of records that fall in A. We prove that the number N_A of records that fall in A is finite a.s. or infinite a.s. according as $H(A)$ is finite or infinite. In the former case we calculate all the moments of N_A and the probability law of N_A. We show that the sequence of record values that fall in A is a Markov chain and give its transition probabilities. Finally, in Section 3, we show that for several sets $A, B, \dots$ that are totally ordered with respect to one another ($A < B$ iff $a < b$ for all $a \in A$, $b \in B$) the restrictions of the records process to the individual sets form independent processes.

As remarked, in Section 4 we show that (A) is satisfied under general topological assumptions and in specific cases.

In Section 5 we return to $\mathbb{R}^2$ and briefly to $\mathbb{R}$. In $\mathbb{R}^2$ we apply the general theory to the case of probability laws in the domain of attraction of a bivariate extreme-value law and calculate hazard measure for laws with independent components.

Finally, in Section 6 we briefly discuss another notion of "record" based on a general partial order.

2. Records in a chosen subset.

We assume (A) *throughout this section.* Define *hazard measure* H on $(S, \mathscr{S})$ by

$$H(A) := \int_{A \cap S_\mu} \{1/\mu((-\infty; x)^c)\} \, d\mu(x),$$

where ^{c} denotes complement in S. So μ and H are mutually absolutely continuous, with Radon–Nikodym derivatives

$$\frac{d\mu}{dH}(x) = \mu\big((-\infty, x)^c\big), \qquad x \in S,$$

$$\frac{dH}{d\mu}(x) = \begin{cases} 1/\mu\big((-\infty, x)^c\big), & x \in S_\mu, \\ 0, & x \notin S_\mu. \end{cases}$$

We now fix a set $A \in \mathscr{S}$ and define the sequence of records that fall in A as follows. Let $L_0^A := 0$, $X_0^A := -\infty$. For $n = 1, 2, 3, \ldots$ define integer-valued r.v.'s

$$L_n^A := \begin{cases} \infty & \text{if } L_{n-1}^A = \infty, \\ \inf\{m \colon m > L_{n-1}^A, X_m \in A, X_m > X_k, k = 0, 1, \ldots, m-1\} & \text{if } L_{n-1}^A < \infty. \end{cases}$$

(Here, $\inf \varnothing := +\infty$.) For $n = 0, 1, 2, \ldots$ let

$$R_n^A := \begin{cases} +\infty & \text{if } L_n^A = \infty, \\ X_{L_n^A} & \text{if } L_n^A < \infty. \end{cases}$$

When $A = S$ we omit the superscript A from these symbols. Thus $\{R_n\}_{n \geq 0}$ is the record-value sequence, which is nondecreasing in S^*, strictly increasing in $\{-\infty\} \cup S$ and starts at $-\infty$. It is immediate that the set of R_n that fall in $A \cup \{-\infty, \infty\}$ gives, after reindexing, the sequence $\{R_n^A\}$.

Let $\mathscr{R} := \{R_n, n \geq 1\} \cap S$ be the set of all records, excluding those in the cemetery states. Our first result concerns $N_A := \sup\{n \colon L_n^A < \infty\} = \operatorname{card}(A \cap \mathscr{R})$, the number of records that fall in A.

THEOREM 2.1. *Let $A \in \mathscr{S}$. Then $P(N_A < \infty) = 1$ or $P(N_A = \infty) = 1$ according as $H(A) < \infty$ or $H(A) = \infty$.*

PROOF. Set

$$A_n := [X_n \in A \cap \mathscr{R}] = [X_n > X_k,\ k = 1, \ldots, n-1,\ X_n \in A].$$

For $m > n$,

$$A_m \subset \big[X_m > X_j,\ j = n+1, \ldots, m-1,\ X_m \in A\big]$$

and the latter event has the same probability as A_{m-n} and is independent of A_n. Hence $P(A_n \cap A_m) \leq P(A_n)P(A_{m-n})$, which is the condition of the Kochen–Stone extension of the Borel–Cantelli lemma [cf. Chow and Teicher (1978), page 101]. Thus $P(A_n \text{ i.o.}) = 0$ iff $\Sigma P(A_n) < \infty$.

Since $[N_A = \infty] = [A_n \text{ i.o.}]$ and

$$\sum_{n=1}^{\infty} P(A_n) = \sum_{n=1}^{\infty} \int_A \mu^{n-1}(-\infty, x)\, d\mu(x)$$

$$= \sum_{n=1}^{\infty} \int_{A \cap S_\mu} \mu^{n-1}(-\infty, x)\, d\mu(x) \quad \text{(from Proposition 1.1)}$$

$$= \int_{A \cap S_\mu} \frac{d\mu(x)}{\mu((-\infty, x)^c)} = H(A),$$

we get $P(N_A = \infty) = 0$ iff $H(A) < \infty$. The rest follows since by the Hewitt–Savage zero–one law $P(A_n \text{ i.o.}) = 0$ or 1. □

The last calculation also shows that $H(A)$ *is the expected number of records in* A. Some explicit calculations of H in case S is $\mathbb{R}$ or $\mathbb{R}^2$ are in Section 5. Here we note only that when A is a singleton, N_A is a 0- or 1-valued r.v., so $H\{x\}$ is the probability that there is a record at x,

$$H\{x\} = P(x \in \mathscr{R}), \qquad x \in S.$$

One should think of $H(dx)$ as the "probability that there is a record at dx."

We now calculate all the moments of N_A. We need the set of strictly ordered n-tuples of elements of A,

$$A^{(n)} := \{(x_1, \dots, x_n) \colon x_1 < x_2 < \cdots < x_n, \text{ all } x_i \in A\},$$

this being a subset of the set of all n-tuples of elements of A,

$$A^n = \{(x_1, \dots, x_n) \colon \text{all } x_i \in A\}.$$

THEOREM 2.2. *If* $H(A) < \infty$ *then, for* $k = 1, 2, \dots,$

$$E(N_A^k) = H(A) + c_k^{(2)} H_2(A) + \cdots + c_k^{(k)} H_k(A) > \infty,$$

where

$$H_j(A) := \int_{A^{(j)}} dH(x_1) \cdots dH(x_j) \tag{2.1}$$

and

$$c_k^{(j)} := \sum_{i=1}^{k} (-)^{j-i} \binom{j}{i} i^k = j!S(k, j), \qquad j = 1, \dots, k, \tag{2.2}$$

$S(k, j)$ *being a Stirling number of the second kind, in the notation of Riordan* (1978).

PROOF. Let $I_m := 1[X_m \in A \cap \mathscr{R}]$. We have

$$N_A^k = (I_1 + I_2 + \cdots)^k. \tag{2.3}$$

The I_m are idempotent: $I_m^2 = I_m$. Hence on expanding the product,

$$(2.4)\quad N_A^k = \sum_n I_n + c_k^{(2)} \sum\sum_{n_1<n_2} I_{n_1} I_{n_2} + \cdots + c_k^{(k)} \sum\sum_{n_1<n_2<\cdots<n_k} I_{n_1} I_{n_2} \cdots I_{n_k},$$

where $c_k^{(j)}$ is the number of ways of choosing one term from each sum in (2.3), such that the set of terms so obtained is $\{I_1, \dots, I_j\}$. Thus $c_k^{(j)}$ is the number of ways of putting k different objects into j different cells, with no cell empty. Its identification with $j!S(k, j)$ is given by Riordan (1978), Chapter 5, Section 5, and the formula (2.2) by Riordan (1978), Chapter 5, Problem 1.

In (2.4) the first term on the right-hand side is N_A, with expectation $H(A)$. We calculate

$$E\Big(\sum\sum_{n_1<\cdots<n_j} I_{n_1} \cdots I_{n_j}\Big)$$

$$= \sum\sum_{n_1<\cdots<n_j} \int_{A^{(j)}} \mu^{n_1-1}(-\infty, x_1)\, d\mu(x_1) \mu^{n_2-n_1-1}(-\infty, x_2)\, d\mu(x_2)$$

$$\times \cdots \times \mu^{n_j-n_{j-1}-1}(-\infty, x_j)\, d\mu(x_j)$$

$$= \int_{A^{(j)}} \sum_{m_1=1}^{\infty} \mu^{m_1-1}(-\infty, x_1)\, d\mu(x_1) \sum_{m_2=1}^{\infty} \mu^{m_2-1}(-\infty, x_2)\, d\mu(x_2)$$

$$\times \cdots \times \sum_{m_j=1}^{\infty} \mu^{m_j-1}(-\infty, x_j)\, d\mu(x_j)$$

(using Fubini's theorem, and taking new summation indices $m_1 := n_1$, $m_2 := n_2 - n_1, \dots, m_j := n_j - n_{j-1}$)

$$= \int_{A^{(j)}} \frac{d\mu(x_1)}{1 - \mu(-\infty, x_1)} \cdots \frac{d\mu(x_j)}{1 - \mu(-\infty, x_j)},$$

which is $H_j(A)$, whence the conclusion. □

COROLLARY 2.3. *If $H(A) < \infty$, then* $\operatorname{var} N_A = H(A)(1 - H(A)) + 2H_2(A)$.

Further results all depend on the quantity

$$Q(A) := P(N_A = 0) = P\big(L_1^A = \infty\big), \qquad A \in \mathscr{S},$$

which we now calculate. Let

$$A_x^y := A \cap (x, y), \qquad x, y \in \mathscr{S}^*,\ x < y;\ A \subseteq S^*.$$

The result below uses the functions H_j defined in (2.1). Note $H_1 \equiv H$.

THEOREM 2.4. *Let $A \in \mathscr{S}$. If $H(A) = \infty$, then $Q(A) = 0$. If $H(A) < \infty$, then*

$$(2.5)\qquad Q(A) = 1 - H_1(A) + H_2(A) - H_3(A) + \cdots.$$

Equivalently, when $H(A) < \infty$, $Q(A)$ *is given by*

$$Q(A) = 1 - \int_A Q(A^x_{-\infty})\, dH(x), \tag{2.6}$$

where $f(x) = Q(A^x_{-\infty})$ *is the unique bounded measurable solution* $f\colon A \to \mathbb{R}$ *of the integral equation*

$$f(y) = 1 - \int_{A^y_{-\infty}} f(x)\, dH(x), \qquad y \in A. \tag{2.7}$$

REMARK. When S is totally ordered, but not in general, (2.5) and (2.7) reduce to standard formulae for product integrals (the Peano series and Volterra integral equation, respectively). See Gill and Johansen (1987).

PROOF. If $H(A) = \infty$ it follows from Theorem 2.1 that $Q(A) = 0$. So now assume $H(A) < \infty$. We first show that Q satisfies (2.6). Let $E_0(x) := \Omega$ and for $n = 1, 2, \dots$ let

$$E_n(x) := \bigcap_{i=1}^{n} \left[X_i \in (-\infty, x) \setminus (\mathscr{R} \cap A^x_{-\infty})\right].$$

Then the event $[L_1^A = n]$ is simply $E_{n-1}(X_n) \cap [X_n \in A]$. So

$$\begin{aligned} 1 - Q(A) &= \sum_{1 \le n < \infty} \int_A P(E_{n-1}(x))\, d\mu(x) \\ &= \sum_{1 \le n < \infty} \int_A P(E_{n-1}(x)) \mu((-\infty, x)^c)\, dH(x) \\ &= \sum_{1 \le n < \infty} \int_A P\big(E_{n-1}(x) \cap \left[X_n \in (-\infty, x)^c\right]\big)\, dH(x) \\ &\qquad\qquad [\text{since } X_n \text{ independent of } E_{n-1}(x)] \\ &= \int_{A \cap S_\mu} \sum_{1 \le n < \infty} P\big(E_{n-1}(x) \cap \left[X_n \in (-\infty, x)^c\right]\big)\, dH(x). \end{aligned}$$

Let $\sigma_x := \inf\{n\colon X_n \in (-\infty, x)^c\}$. Then $E_{n-1}(x) \cap [X_n \in (-\infty, x)^c]$ is the event $\sigma_x = n] \cap \bigcap_{i=1}^{n-1}[X_i \notin \mathscr{R} \cap A^x_{-\infty}]$. For $x \in S_\mu$, σ_x is a.s. finite, so the probability of the union of these events is $Q(A^x_{-\infty})$. Thus the last integral above equals $\int_A Q(A^x_{-\infty})\, dH(x)$, whence (2.6).

We may replace A here by $A^y_{-\infty}$; hence $Q(A^y_{-\infty})$ satisfies (2.7). Let us substitute this equation repeatedly into (2.6). By induction we obtain that for each n,

$$\begin{aligned} Q(A) = 1 - H_1(A) + H_2(A) - \cdots + (-)^{n-1} H_{n-1}(A) \\ + (-)^n \int_{A^{(n)}} Q(A^{x_n}_{-\infty})\, dH(x_n) \cdots dH(x_1). \end{aligned}$$

So for (2.5) we need the last term to tend to 0. In absolute value it is at most

$$\int_{A^{(n)}} dH(x_n) \cdots dH(x_1)$$

$$= (1/n!) \int_{A^n \cap \{(x_1, \ldots, x_n):\, x_1, \ldots, x_n \text{ are totally ordered}\}} dH(x_n) \cdots dH(x_1)$$

$$\leq (1/n!) \int_{A^n} dH(x_n) \cdots dH(x_1) = (H(A))^n / n! \to 0,$$

establishing (2.5).

Finally, we show the solution to (2.7) is unique. Let f be any bounded measurable solution. Then, iterating as above,

$$f(y) = 1 - H_1(A^y_{-\infty}) + H_2(A^y_{-\infty}) + \cdots + (-)^{n-1} H_{n-1}(A^y_{-\infty})$$
$$+ (-)^n \int_{(A^y_{-\infty})^{(n)}} f(x_n)\, dH(x_n) \cdots dH(x_1).$$

The last term tends to 0, as in the calculation above. So

$$f(y) = 1 - H_1(A^y_{-\infty}) + H_2(A^y_{-\infty}) - H_3(A^y_{-\infty}) + \cdots$$

and is thus unique. □

We remark that while $Q(A) = 0$ is forced by $H(A) = \infty$, it can occur even though $H(A) > \infty$. For instance, $Q(S) = 0$ whatever the value of $H(S)$, as the first observation is always a record.

For one class of sets Q has a simple formula. The set $A \subseteq S$ is called a *lower set* if $\downarrow A = A$, where

$$\downarrow A := \{ y \in S\colon\, y \leq x \text{ for some } x \in A \},$$

Corollary 2.5. *If A is a lower set, then $Q(A) = \mu(A^c)$.*

Proof. Immediate, because $N_A > 0$ iff $X_1 \in A$. It also follows from (2.6), since $A^x_{-\infty} = (-\infty, x)$ when $x \in A$. □

We use Q to characterize $\{R_n\}$ as a Markov chain. We also have a little to say about the times between records in A,

$$\Delta_n^A := \begin{cases} \infty & \text{if } L_{n+1}^A = \infty, \\ L_{n+1}^A - L_n^A & \text{if } L_{n+1}^A < \infty, \end{cases} \qquad n = 0, 1, 2, \ldots.$$

For $n = 1, 2, \ldots$ let $\mathscr{A}_n^A := \sigma(X_i, i \leq L_n^A)$, that is, $\mathscr{A}_n^A$ is the σ-algebra of all events E such that $E \cap [L_n^A = k] \in \sigma(X_1, \ldots, X_k)$ for all $k = 1, 2, \ldots$. Let $\mathscr{A}_0^A := \{\varnothing, \Omega\}$. Recall that the deterministic behaviour of $\{R_n^A\}_{0 \leq n < \infty}$ is that it has initial state $-\infty$, is strictly increasing until it reaches the absorbing state

$+\infty$, and whether or not it reaches that state is determined by finiteness or not of $H(A)$.

THEOREM 2.6. *$\{R_n^A, \mathscr{A}_n^A\}_{0 \le n < \infty}$ is a Markov chain with stationary transition laws*

$$P\big(R_n^A \in dy | R_{n-1}^A = x\big) = 1[y \in A_x^\infty] Q(A_x^y)\, dH(y), \qquad y \in S, \tag{2.8}$$

$$P\big(R_n^A = \infty | R_{n-1}^A = x\big) = Q(A_x^\infty), \qquad x \in S^*,\ n = 1, 2, \dots. \tag{2.9}$$

Given the random set $\mathscr{R} \cap A$, the Δ_n^A are conditionally independent, with conditional laws

$$P\big(\Delta_n^A = k | \mathscr{R} \cap A\big) = P\big(\Delta_n^A = k | R_n^A, R_{n+1}^A\big) = g_k^A\big(R_n^A, R_{n+1}^A\big) \quad a.s., \tag{2.10}$$
$$n = 0, 1, 2, \dots,\ k = 0, 1, 2, \dots, \infty,$$

where, for all $k \in \mathbb{N}$,

$$g_k^A(x, y) := \begin{cases} q_k^A(x, y)/Q(A_x^y), & x < y,\ y \in S_\mu, \\ 0, & \text{other } (x, y) \in (S^*)^2, \end{cases} \tag{2.11}$$

and, for $-\infty \le x < y < \infty$, $k \in \mathbb{N}$,

$$q_k^A(x, y) := P\big(X_k \text{ is the first observation in } (-\infty, y)^c, \text{ and none of } X_1, \dots, X_{k-1} \text{ is in } \mathscr{R} \cap A_x^y\big).$$

NOTE. In (2.11) the denominator $Q(A_x^y)$ is positive where it appears, because it is at least the probability that the first observation in $(-\infty, x)^c$ is also in $(-\infty, y)^c$, that is, at least $\mu\big((-\infty, y)^c\big)/\mu\big((-\infty, x)^c\big) > 0$.

PROOF. The idea is a form of spatial Markov property. Let $\mathscr{A}_n^{A-} := \sigma(X_i,\ i < L_n^A)$. We have that

$$\text{the set } \{\Delta_j^A,\ j \ge n,\ R_j^A,\ j > n\} \text{ is conditionally independent of } \mathscr{A}_n^{A-}, \text{ given } R_n^A. \tag{2.12}$$

For when $L_n^A = m < \infty$, we necessarily have $x_j < R_n^A$ for all $j < m$, and no further information about $X_1, \dots, X_{m-1}$ is needed to determine the set of r.v.'s mentioned in (2.12).

From (2.12) the Markov property of $\{R_n^A\}$ follows, and, furthermore, the conditional independence of the Δ_n^A and the left-hand equality in (2.10). It remains to evaluate the various conditional laws.

We first prove that, for $m, n \geq 0,\ k \geq 1,\ y \in S$,

$$P\left(\Delta_n^A = k,\ R_{n+1}^A \in dy \mid X_m\right) = 1\left[y \in A_{X_m}^{\infty}\right] q_k^A(X_m, y)\, dH(y) \tag{2.13}$$

a.s. on $\left[L_n^A = m\right]$.

For take any $B, C \in \mathscr{S}^*$ and let $B' \in (\mathscr{S}^*)^m$ be such that

$$[(X_1, \dots, X_m) \in B'] = \left[L_n^A = m,\ X_m \in B\right].$$

Then, conditioning on $X_1 = x_1, \dots, X_m = x_m,\ X_{m+k} = y$,

$$\begin{aligned}
&P\left(X_m \in B,\ L_n^A = m,\ \Delta_n^A = k,\ R_{n+1}^A \in C\right)\\
&= \int_{B'}\int_{C \cap A_{x_m}^{\infty} \cap S_\mu} P\Bigg(X_{m+i} \in ((-\infty, y) \setminus A)\\
&\qquad \cup \left(A_{-\infty}^{y} \setminus \left((x_m, \infty) \cap \bigcap_{m+1}^{m+i-1} (X_j, \infty)\right)\right),\ i = 1, \dots, k-1\Bigg)\\
&\qquad \times d\mu(y)\, d\mu(x_1) \cdots d\mu(x_m)\\
&= \int_{B'}\int_{C \cap A_{x_m}^{\infty} \cap S_\mu} P\Bigg(X_{m+i} \in ((-\infty, y) \setminus A)\\
&\qquad \cup \left(A_{-\infty}^{y} \setminus \left((x_m, \infty) \cap \bigcap_{m+1}^{m+i-1} (X_j, \infty)\right)\right),\ i = 1, \dots, k-1,\\
&\qquad X_{m+k} \in (-\infty, y)^c\Bigg)\, dH(y)\, d\mu(x_1) \cdots d\mu(x_m)\\
&= \int_{B'}\int_{C} 1\left[y \in A_{x_m}^{\infty}\right] q_k^A(x_m, y)\, dH(y)\, d\mu(x_1) \cdots d\mu(x_m)
\end{aligned}$$

and this establishes (2.13). From (2.13), immediately,

$$P\left(\Delta_n^A = k,\ R_{n+1}^A \in dy \mid R_n^A\right) = 1\left[y \in A_{R_n^A}^{\infty}\right] q_k^A\left(R_n^A, y\right) dH(y) \quad \text{a.s.} \tag{2.14}$$

We claim that

$$\sum_{k=1}^{\infty} q_k^A(x, y) = Q(A_x^y) \quad \text{if } y \in S_\mu. \tag{2.15}$$

For if $\mu((-\infty, y)^c) > 0$, then there certainly will be a first observation in $(-\infty, y)^c$, and the left-hand side is the probability that no prior observation is in $\mathscr{R} \cap A_x^y$, hence is the probability that $\mathscr{R} \cap A_x^y = \varnothing$.

Now sum k in (2.15). Then

$$P\left(R_{n+1}^A \in dy \mid \mathscr{A}_n^A\right) = 1\left[y \in A_{R_n^A}^{\infty}\right] Q\left(A_{R_n^A}^{y}\right) dH(y) \quad \text{a.s.,}$$

because $dH(y) = 0$ if $y \notin S_\mu$ and otherwise (2.14) applies. Thus we have proved (2.8). From (2.6)

$$Q(A_x^\infty) = 1 - \int_{A \cap (x, \infty)} Q(A_x^y)\, dH(y);$$

then (2.9) follows.

Lastly, the second equality in (2.10) follows from (2.14) and (2.8). □

The above result retrieves a small fraction of the rich structure of records in a totally ordered set—say $\mathbb{R}$. In $\mathbb{R}$ the Markov chain $\{R_n\}$ is the sequence of points of a completely random point process [Shorrock (1972, 1974) and Goldie and Rogers (1984)], specializing to a Poisson process if μ is continuous. And the conditionally independent interrecord times Δ_n have conditional laws (2.10) which depend only on R_n, not also on R_{n+1}, and are in fact geometric distributions.

COROLLARY 2.7. *If $H(A) < \infty$, then the law of N_A is $P(N_A = 0) = Q(A)$,*

$$P(N_A = j)$$

$$= \int_{A^{(j)}} Q(A_{-\infty}^{x_1}) Q(A_{x_1}^{x_2}) \cdots Q(A_{x_{j-1}}^{x_j}) Q(A_{x_j}^{\infty})\, dH(x_1) \cdots dH(x_j), \tag{2.16}$$

$$j = 1, 2, \ldots .$$

Also

$$P(N_A \geq j)$$

$$= \int_{A^{(j)}} Q(A_{-\infty}^{x_1}) Q(A_{x_1}^{x_2}) \cdots Q(A_{x_{j-1}}^{x_j})\, dH(x_1) \cdots dH(x_j), \tag{2.17}$$

$$j = 1, 2, \ldots .$$

And N_A has an entire characteristic function (moment generating function finite everywhere).

PROOF. $P(N_A = j) = P(R_j^A \in A,\ R_{j+1}^A = \infty)$, and (2.16) follows by conditioning successively on $\mathscr{A}_j^A, \mathscr{A}_{j-1}^A, \ldots, \mathscr{A}_1^A$ and employing (2.8) and (2.9). Similarly for $P(N_A \geq j) = P(R_j^A \in A)$.

Finally, $P(N_A = j) \leq \int_{A^{(j)}} dH(x_1) \cdots dH(x_j) \leq (H(A))^j / j!$ as noted in the proof of Theorem 2.5. So $Ee^{\lambda N_A} < \infty$ for all real λ. □

Again, when $S = \mathbb{R}$, N_A has a Poisson law if μ is continuous. Here, for general S, we have only a Poisson bound: As the last calculation above shows, for each j, $P(N_A = j) \leq e^{H(A)} p_j$, where $\{p_j\}$ is the Poisson law of parameter $H(A)$.

3. Records in several subsets. *We assume* (A) *throughout this section.* We investigate the relationships between the restrictions of the record-value process to several subsets of S.

THEOREM 3.1. *Let $A, B \subseteq S$. If $H(A \cup B) < \infty$, then*

$$\operatorname{cov}(N_A, N_B) = H(A \cap B) - H(A)H(B) + \int_A \int_{B \cap (x,\infty)} dH(y)\, dH(x)$$
$$+ \int_B \int_{A \cap (x,\infty)} dH(y)\, dH(x).$$

PROOF.

$$\begin{aligned} N_A N_B &= \sum_{m=1}^{\infty} 1[X_m \in A \cap \mathscr{R}] \sum_{n=1}^{\infty} 1[X_n \in B \cap \mathscr{R}] \\ &= N_{A \cap B} + \sum_{m<n}\sum 1[X_m \in A \cap \mathscr{R},\, X_n \in B \cap \mathscr{R}] \\ &\quad + \sum_{n<m}\sum 1[X_m \in A \cap \mathscr{R},\, X_n \in B \cap \mathscr{R}]. \end{aligned}$$

The first term on the right has expectation $H(A \cap B)$. The second has expectation

$$\begin{aligned} &\sum_{m<n}\sum \int_A \int_{B \cap (x,\infty)} \mu^{n-m-1}(-\infty, y)\, d\mu(y)\, \mu^{m-1}(-\infty, x)\, d\mu(x) \\ &= \int_A \int_{B \cap (x,\infty)} \sum_{l=1}^{\infty} \mu^{l-1}(-\infty, y)\, d\mu(y) \sum_{m=1}^{\infty} \mu^{m-1}(-\infty, x)\, d\mu(x) \\ &= \int_A \int_{B \cap (x,\infty)} dH(y)\, dH(x). \end{aligned}$$

Similarly for the third term. The result follows. □

It is clear from this result that in general we lose the property of *complete randomness* of the record-value process in $\mathbb{R}$: That N_{A_i} for disjoint A_i are independent r.v.'s. But we can rescue it partially by specializing to sets A_i that are *totally ordered* with respect to each other. For subsets A, B of S^* we write

$$A < B \quad \text{when } a < b \text{ for all } a \in A,\ b \in B.$$

Note that if $A < B$, then $\operatorname{cov}(N_A, N_B) = 0$. In fact more is true.

THEOREM 3.2. *Let $\{A(i)\}_{i \in I}$ be a collection of measurable subsets of S, indexed by some finite or infinite set of integers I that are totally ordered with respect to one another,*

$$A(i) < A(j) \quad \textit{whenever } i < j, \qquad i \in I,\ j \in I.$$

Then the processes $\{R_n^{A(i)}\}_{0 \le n < \infty}$, $i \in I$, are independent.

PROOF. It suffices to consider finite I. Since $H(A) = EN_A$, we have $H(A) = 0$ iff $N_A = 0$ a.s., and we may remove any $A(i)$ that fall in this trivial class. We consider only the case $I = \{1, 2\}$, because the case of general finite I is a straightforward elaboration. Pick $A, B \in \mathcal{S}$ with $A < B$. It suffices to check that for $k, l \in \mathbb{N}$, and measurable subsets $E \subseteq A^{(k)}$, $F \subseteq B^{(l)}$, the events

$$(3.1) \qquad U := \left[(R_1^A, \dots, R_k^A) \in E \right], \qquad V := (R_1^B, \dots, R_l^B) \in F \right]$$

are independent. For then the similar statements involving some of the R_j^A, R_j^B infinite follow by complementations.

The first event U belongs to the pre-L_k^A field $\mathscr{A}_k^A$. In (3.1) V will be independent of U provided V belongs to the post-L_k^A field $\sigma(X_\iota, i > L_k^A)$. [For the independence of the pre- and post-σ fields, for any stopping time σ, see, e.g., Chung (1974), Exercise 10, Section 8.2.]

Now when $L_k^A = m < \infty$ we have $X_i < X_m \in A$ for $i < m$, so for any X_{m-n}, $n \geq 1$, that falls in B we automatically have $X_i < X_{m-n}$ for all $i \leq m$. Thus to check whether $X_{m-n} \in \mathscr{R}$ we need look only at the r.v.'s $X_{m-1}, \dots, X_{m-n}$. Hence V indeed belongs to the post-L_k^A field. □

COROLLARY 3.3. *For $\{A(i)\}$ as in Theorem 3.2, the r.v.'s $N_{A(\iota)}$, $i \in I$, are independent.*

COROLLARY 3.4. *Suppose I is finite or countable and that the points $\{x_i\}_{\iota \in I}$ are totally ordered in S. Then the events $[x_i \in \mathscr{R}]$ are independent.*

REMARK. The reasoning of Theorem 3.2 may be applied to other orderings. For example, in $\mathbb{R}^2$ consider ordering convex hulls in the following manner: Let $\mathscr{C}(x_1, \dots, x_n)$ denote the convex hull of the points $x_1, \dots, x_n \in \mathbb{R}^2$. For two polygons C_1, C_2 in $\mathbb{R}^2$ say $C_1 < C_2$ if C_1 is contained in the region enclosed by C_2 and the vertices of C_2 are disjoint from those of C_1. Let $X_1, X_2, \dots$ be i.i.d. random vectors in $\mathbb{R}^2$. Define record times recursively as follows: Let $\mathscr{C}_n := \mathscr{C}(X_1, \dots, X_n)$ and set $L(1) := 3$ and

$$L(n+1) := \inf\{k > L(n) \colon \mathscr{C}_k > \mathscr{C}_{L(n)}\}.$$

The set of record values is $\mathscr{R} := \{\mathscr{C}_{L(n)}\}_{n \geq 1, L(n) < \infty}$. Then, for instance, there is a result like the last corollary above: If $C_1 < C_2$ are convex polygons we have

$$[C_1 \in \mathscr{R}] \text{ and } [C_2 \in \mathscr{R}] \text{ independent.}$$

The proof is like that for Theorem 3.2: Set $\sigma_\iota := \inf\{L(k) \colon \mathscr{C}_{L(k)} = C_\iota\}$ so that

$$\begin{aligned} P(\sigma_1 < \infty, \sigma_2 < \infty) = \sum_{1 \leq m < \infty} & P(\sigma_1 = m, \\ & m + \inf\{n \colon \mathscr{C}(X_{m-1}, \dots, X_{m-n}) = C_2\} < \infty) \\ = \sum_{1 \leq m < \infty} & P(\sigma_1 = m) P(\sigma_2 < \infty) = P(C_1 \in \mathscr{R}) P(C_2 \in \mathscr{R}). \end{aligned}$$

4. Topology. Our aim here is to show that assumption (A), and hence the theorems of Sections 2 and 3, hold for familiar topological spaces possessing a partial order, including those listed as (a)–(h) in Section 1.

PROPOSITION 4.1. *Let S be a topological space having a countable base for its topology. Give $S \times S$ the product topology. Suppose S has a* closed order, *that is, a partial order such that the graph $G_{\leq}$ is closed in $S \times S$. The* (A) *holds.*

PROOF. Existence of the countable base makes the Borel σ-algebra in $S \times S$ coincide with the product σ-algebra $\mathcal{S} \times \mathcal{S} = \mathcal{B}(S) \times \mathcal{B}(S)$ [cf. Dellacherie and Meyer (1978), I.8(c)], so that closed sets in $S \times S$ are product-measurable. Thus $G_{\leq}$ is measurable, hence so is $G_{<}$ as remarked after Proposition 1.1. □

The closed-order assumption is also known as "order-Hausdorff" and implies that S is Hausdorff [Nachbin (1965), Proposition 1].

COROLLARY 4.2. *Let S be a Hausdorff space having a countable base and suppose that S is also a topological semilattice, that is, a lattice such that the meet operation $(x, y) \to x \wedge y$ is continuous from $S \times S$ into S. Then* (A) *holds.*

PROOF. The Hausdorff property and meet-continuity imply that S has closed order [Gierz, Hofmann, Keimel, Lawson, Mislove and Scott (1980), VI, Proposition 1.14]. □

DEFINITION. Let S be a set of real-valued functions on a set E. The usual weak order on $\mathbb{R}$ induces a partial order on S, the *induced weak order* $\leq$, by

$$x \leq y \quad \text{iff } x(e) \leq y(e), \quad \text{all } e \in E.$$

This is the partial order employed in examples (b), (d), (e) and (g) in Section 1. Of these, (b) ($\mathbb{R}^d$ with induced weak order) satisfies assumption (A) trivially, as does (d) ($\mathbb{R}^{\mathbb{N}}$ with product topology and induced weak order) by Proposition 4.1. In case (e) ($C[0, 1]$ with the topology of uniform convergence and induced weak order) it is easy to check $G_{\leq}$ is closed, so Proposition 4.1 applies. The same holds for $C(E)$ where E is compact Hausdorff with a countable base, and for $C[0, \infty]$ with the topology of locally uniform convergence. Two final easy examples are (a) and (c) ($\mathbb{R}^d$ and $\mathbb{R}^{\mathbb{N}}$ with induced strict order) where we have (A) at once because $G_{<}$ is open and so product-measurable.

An interesting extension of (b) results from defining $x \leq y$ to mean $x - y \in C$, where C is a fixed measurable set in $\mathbb{R}^d$ with the properties

$$C \cap (-C) = \{0\},$$
$$C + C := \{x + y \colon x \in C,\ y \in C\} \subseteq C.$$

For instance, in $\mathbb{R}^2$, C could be a cone $\{(x, y) \colon |y| \leq \theta x\}$ where $\theta > 0$. Example (b) is the special case when C is the nonnegative orthant. Other orthants can equally be used.

For more general examples a framework may be built from *continuous lattices* [Gierz et al. (1980)], but with the order reversed, as in Gerritse (1985). Suppose $(S, \leq)$ is a complete lattice. A subset B of S is *filtered* if every finite subset of B has a lower bound in B. The element $x \in S$ is *way above* $y \in S$, in symbols

$x \gg y$, if for each filtered $B \subseteq S$, with $\inf B \leq y$, there exists $z \in B$ such that $z \leq x$. The complete lattice S is *upper-continuous* if $x = \inf\{y \in S\colon y \gg x\}$ for all $x \in S$. Supposing S to be an upper-continuous lattice, its *upper Lawson topology* is that with sub-base the sets

$$\{y \in S\colon x \gg y\}, \qquad \{y \in S\colon y \leq x\}^c$$

for all $x \in S$.

PROPOSITION 4.3. *Let S be an upper-continuous lattice with its upper Lawson topology; suppose this has countable base. Then* (A) *holds.*

PROOF. $G_{\leq}$ has closed order [Gerritse (1985), Theorem 6.6], so Proposition 4.1 applies. □

Consider now $US(E)$, the space of usc functions on E [example (g) in Section 1]. Under the induced weak order $US(E)$ is an upper-continuous lattice if E is locally quasicompact [Gerritse (1985), Theorem 8.4]. Its upper Lawson topology turns out [Vervaat (1988b), Section 3.9] to be the sup vague topology, with sub-base consisting of all sets

$$\left\{f \in US(E)\colon \sup_{t \in K} f(t) < x\right\}, \qquad \left\{f \in US(E)\colon \sup_{t \in G} f(t) > x\right\}$$

for K quasicompact in E, G open in E and $x \in [-\infty, \infty]$. If E has also a countable base, then it suffices to let G and K run over countable collections of sets [Gerritse (1985), Theorem 5.5; this is obvious for G but not for K]. So $US(E)$ then has countable base, and we obtain the following corollary.

COROLLARY 4.4. *Let E be lqccb and let $S = US(E)$ with the induced weak order and sup vague topology. Then* (A) *holds.*

For example (f), the class $\mathscr{F}(E)$ of closed subsets of E, the sup vague topology of Gerritse (1985), Section 3, is the "Fell" topology [Norberg (1986)] which has sub-base

$$\{F\colon F \cap K = \varnothing\}, \qquad \{F\colon F \cap G \neq \varnothing\}$$

for all quasicompact K, open G. This is the topology used by Matheron (1975); see Vervaat (1988a) for further connections. When E is compact Hausdorff this topology is metrised by Hausdorff distance. Now the map $F \mapsto 1_F$ embeds $\mathscr{F}(E)$ as a closed subset of $US(E)$ [Vervaat (1988b), Theorem 4.1]; also $\mathscr{F}(E)$ as a latttice embeds into the lattice $US(E)$. We obtain the following corollary.

COROLLARY 4.5. *Let E be lqccb. Let $S = \mathscr{F}(E)$ be partially ordered by set-inclusion and have the Fell topology. Then* (A) *holds.*

Finally, consider example (h), the space $\mathscr{P}$ of probability measures on $\mathbb{R}$. Under the topology of narrow ("weak") convergence, every evaluation map

$\mu \mapsto \mu(B)$, where B is a Borel set in $\mathbb{R}$, is measurable. Hence, on $\mathscr{P} \times \mathscr{P}$, every evaluation map $(\mu, \nu) \mapsto (\mu \times \nu)(B)$, where B is a Borel set in $\mathbb{R}^2$, is product-measurable. Now one partial order on $\mathscr{P}$ is: $\mu \leq_{s2} \nu$ iff

$$(4.1) \qquad \begin{aligned} &(\mu \times \nu)\{(x, y): a \leq x < b,\, x \leq y\} \\ &\quad \geq (\mu \times \nu)\{(x, y): a \leq y < b,\, y \leq x\} \end{aligned}$$

for all real $a > b$. Zijlstra and de Kroon (1981) show this is indeed a partial order and give applications. In (4.1) it suffices to let a and b run through the rationals. So this partial order is defined by a countable infinity of inequalities on evaluation maps, hence its graph $G_{\leq}$ is measurable in $\mathscr{P} \times \mathscr{P}$.

For other partial orderings of probability laws (stochastic orderings of r.v.'s) see, for example, Kamae, Krengel and O'Brien (1977), Stoyan and Stoyan (1980) and Ahmed, Alzaid, Bartoszewicz and Kochar (1986). Similar considerations will establish the measurability of $G_{\leq}$.

5. More in $\mathbb{R}$ and $\mathbb{R}^2$. In Shorrock (1970), Theorem 1.2, it is shown that in $\mathbb{R}$ there are finitely many strict records, a.s., if μ has an atom at the supremum x_+ of its support, and infinitely many, a.s., when μ has no atom at x_+. (When $x_+ = \infty$, the latter must occur.) That this follows from Theorem 2.2 is a consequence of the following analytical lemma.

LEMMA 5.1 [Shanbhag (1979), Theorem 2]. *$H(\mathbb{R}) < \infty$ if and only if $\mu\{x_+\} > 0$.*

We turn to $\mathbb{R}^2$. Let

$$\phi(t) := \sum_{k=1}^{\infty} \frac{t^k}{k^2} = \int_0^t \log\left(\frac{1}{1-u}\right)\frac{du}{u}, \qquad 0 \leq t < 1.$$

PROPOSITION 5.2. *If μ on $\mathbb{R}^2$ is a product measure with continuous components, then*

$$H(-\infty, \mathbf{x}] = \phi(\mu(-\infty, \mathbf{x}]), \qquad \mathbf{x} \in \mathbb{R}^2.$$

PROOF. Write $\mu(-\infty, \mathbf{x}] = F(x^{(1)})G(x^{(2)})$ where F, G are continuous d.f.'s (distribution functions) on $\mathbb{R}$. Then

$$\begin{aligned} H(-\infty, \mathbf{x}] &= \int_{-\infty}^{x^{(1)}} \int_{\infty}^{x^{(2)}} \frac{dG(v)\, dF(u)}{1 - F(u)G(v)} \\ &= \int_{-\infty}^{x^{(1)}} \{-\log(1 - F(u)G(x^{(2)}))\} \frac{dF(u)}{F(u)} \\ &= \int_0^{F(x^{(1)})} \{-\log(1 - tG(x^{(2)}))\} \frac{dt}{t} \\ &= \phi(F(x^{(1)})G(x^{(2)})). \end{aligned}$$

□

We remark that when μ on $\mathbb{R}^d$ is a product measure, under the strict order (a) the event that $X_n = (X_n^{(1)}, \dots, X_n^{(d)})$ is a record is simply the intersection of the independent events $[X_n^{(j)}$ is a record among $X_1^{(j)}, \dots, X_n^{(j)}]$. Various one-dimensional results may then be applied. For instance, the probability of each of the latter events is $1/n$; hence

$$P(X_n \in \mathscr{R}) = 1/n^d,$$

and the total number of records is a.s. finite for all $d > 1$.

We now present a class of examples in $\mathbb{R}^2$, which includes the bivariate Cauchy law, where $P(N = \infty) = 1$. Suppose that the bivariate d.f. $F(\mathbf{x}) := \mu(-\infty, \mathbf{x}]$ is continuous and in the domain of attraction of a bivariate extreme-value d.f. G. Give $\mathbb{R}^2$ the strict partial order of (4.1). The number of records in $\{(X_n^{(1)}, X_n^{(2)}),\ n \ge 1\}$ is the same as in $\{(U_1(X_n^{(1)}), U_2(X_n^{(2)})),\ n \ge 1\}$ provided the functions $U_i\colon \mathbb{R} \to \mathbb{R}$ are strictly increasing. Therefore there is no loss of generality in supposing the marginal d.f.'s of G are each equal to $\exp\{-x^{-1}\}$, for $x > 0$, and that the marginal distributions of F are tail equivalent so that $P(X_1^{(1)} > x) \sim P(X_1^{(2)} > x)$ as $x \to \infty$. Assuming F is in the domain of attraction of G leads to F having a tail $1 - F(\mathbf{x})$ which is regularly varying at ∞ with limit measure ν. This means ν concentrates on the punctured compact set $[\mathbf{0}, \boldsymbol{\infty}] \setminus \{\mathbf{0}\}$, topologized so that compact sets are closed sets bounded away from $\mathbf{0}$, and for $\mathbf{x} \in [\mathbf{0}, \boldsymbol{\infty}) \setminus \{\mathbf{0}\}$,

$$\lim_{t\to\infty} (1 - F(t\mathbf{x}))/(1 - F(t\mathbf{1})) = \nu([0, \mathbf{x}]^c).$$

There exists a finite measure S on $\mathscr{B}([0, \pi/2])$ such that

$$\nu\{\mathbf{x}\colon \|\mathbf{x}\| > r,\ \theta(x) \in (\theta_1, \theta_2]\} = r^{-\alpha} S(\theta_1, \theta_2],$$

where $(\|x\|, \theta(x))$ are the polar coordinates of x and $0 \le \theta_1 < \theta_2 \le \pi/2$. Note G is a product measure iff S concentrates on $\{0, \pi/2\}$ [cf. de Haan and Resnick (1977) and Resnick (1987)].

First assume S does not concentrate on $\{0, \pi/2\}$. For such a distribution we show $H(\mathbb{R}^2) = \infty$. We have that for $t > 0$, $M > 1$,

$$\begin{aligned} H(\mathbb{R}^2) &\ge \iint_{\{\mathbf{x}\ge 0\colon t<\|\mathbf{x}\|\le tM\}} \frac{F(d\mathbf{x})}{F((-\infty, \mathbf{x})^c)} \\ &= \iint_{\{\mathbf{x}\ge 0\colon 1<\|\mathbf{x}\|\le M\}} \frac{F(t\,d\mathbf{x})}{F((-\infty, t\mathbf{x})^c)} \\ &\to \iint_{\{\mathbf{x}\ge 0\colon 1<\|\mathbf{x}\|\le M\}} \frac{\nu(d\mathbf{x})}{\nu([0, \mathbf{x})^c)} \end{aligned}$$

as $t \to \infty$. Now this last integral exceeds

$$\iint_{\{\mathbf{x}\geq\mathbf{0}:\,1<\|\mathbf{x}\|\leq M\}} \frac{\nu(d\mathbf{x})}{\nu\{\mathbf{y}\geq\mathbf{0}:\,\|\mathbf{y}\|>x^{(1)}\wedge x^{(2)}\}}$$

$$= \iint_{\{\mathbf{x}\geq\mathbf{0}:\,1<\|\mathbf{x}\|\leq M,\,\theta(\mathbf{x})\in[0,\pi/4)\}} \frac{\nu(d\mathbf{x})}{\nu\{\mathbf{y}\geq\mathbf{0}:\,\|\mathbf{y}\|>x^{(2)}\}}$$

$$+ \iint_{\{\mathbf{x}\geq\mathbf{0}:\,1<\|\mathbf{x}\|\leq M,\,\theta(\mathbf{x})\in[\pi/4,\pi/2]\}} \frac{\nu(d\mathbf{x})}{\nu\{\mathbf{y}\geq\mathbf{0}:\,\|\mathbf{y}\|>x^{(1)}\}}$$

$$= \iint_{1<\|\mathbf{x}\|\leq M,\,\theta(\mathbf{x})\in[0,\pi/4)} \frac{(x^{(2)})^{\alpha}}{S[0,\pi/2)}\nu(d\mathbf{x})$$

$$+ \iint_{1<\|\mathbf{x}\|\leq M,\,\theta(\mathbf{x})\in[\pi/4,\pi/2]} \frac{(x^{(1)})^{\alpha}}{S[0,\pi/2)}\nu(d\mathbf{x})$$

$$= S^{-1}[0,\pi/2)\int_1^M r^{\alpha}\alpha r^{-\alpha-1}\,dr\left\{\int_0^{\pi/4}(\sin\theta)^{\alpha}S(d\theta)+\int_{\pi/4}^{\pi/2}(\cos\theta)^{\alpha}S(d\theta)\right\}$$

$$= S^{-1}[0,\pi/2)\alpha\log M\left\{\int_0^{\pi/4}(\sin\theta)^{\alpha}S(d\theta)+\int_{\pi/4}^{\pi/2}(\cos\theta)^{\alpha}S(d\theta)\right\}.$$

Since M is at our disposal, we may let $M \to \infty$. Provided the term in the braces is positive, this shows $H(\mathbb{R}^2) = \infty$. The term in the braces is 0 iff S concentrates on $\{0, \pi/2\}$, so this case must be excluded. Concentration on $\{0, \pi/2\}$ corresponds to the limiting extreme-value distribution being a product measure.

In fact, when S concentrates on $\{0, \pi/2\}$ we may show $H(\mathbb{R}^2) < \infty$ as follows. Write, for $t > 0$,

$$\iint_{\mathbb{R}^2} = \iint_{\{\mathbf{x}:\,x^{(1)}\vee x^{(2)}\leq t\}} + \iint_{\{\mathbf{x}:\,x^{(1)}\vee x^{(2)}>t\}} = \mathrm{I}+\mathrm{II}.$$

The integrand is $\{1/F((-\infty,\mathbf{x})^c)\}F(d\mathbf{x})$. Now with $\mathbf{1} = (1,1)$,

$$\mathrm{I} \leq F((-\infty,t\mathbf{1}])/F\big((-\infty,t\mathbf{1})^c\big) < \infty.$$

For II we write

$$\mathrm{II} = \iint_{\{\mathbf{x}:\,x^{(1)}>t,\,x^{(2)}\leq t\}} + \iint_{\{\mathbf{x}:\,x^{(1)}\leq t,\,x^{(2)}>t\}} + \iint_{\{\mathbf{x}:\,x^{(1)}\wedge x^{(2)}>t\}}$$

$$= \mathrm{IIa}+\mathrm{IIb}+\mathrm{IIc}.$$

For IIc we have

$$\mathrm{IIc} = \iint_{(1,\infty]^2} F(t\,d\mathbf{x})/F\big((-\infty,t\mathbf{x})^c\big)$$

$$\to \iint_{(1,\infty]^2} \nu(d\mathbf{x})/\nu\big([0,\mathbf{x})^c\big) = 0$$

since ν concentrates on the coordinate axes through $\mathbf{0}$. For IIa we get

$$\begin{aligned}\text{IIa} &= \iint_{[x^{(1)} \le 1,\, x^{(2)} > 1]} F(t\,d\mathbf{x}) / P\big(X_1^{(1)} \ge tx \text{ or } X_1^{(2)} \ge tx^{(2)}\big) \\ &\le \iint_{[x^{(1)} \le 1,\, x^{(2)} > 1]} F(t\,d\mathbf{x}) / P\big(X_1^{(1)} \ge t \text{ or } X_1^{(2)} \ge tx^{(2)}\big) \\ &\to \iint_{[x^{(1)} \le 1,\, x^{(2)} > 1]} \nu(d\mathbf{x}) / \nu[\mathbf{0}, (1, x^{(2)}))\end{aligned}$$

and since ν has no mass on $(\mathbf{0}, \boldsymbol{\infty})$, the denominator equals $S\{0\} + S\{\pi/2\}/x^{(2)}$ and therefore

$$\begin{aligned}\limsup_{t \to \infty} \text{IIa} &\le \int_1^\infty \frac{S\{\pi/2\}\, d(-y^{-1})}{S\{0\} + S\{\pi/2\} y^{-1}} \\ &= S\{\pi/2\}^{-1} \log(1 + S\{\pi/2\}/S\{0\}) \\ &< \infty.\end{aligned}$$

We may handle IIb similarly and we see that for some t sufficiently large, $\text{II} < \infty$.

We summarize this discussion.

THEOREM 5.3. *Suppose F is continuous and in the domain of attraction of the bivariate extreme-value distribution G. Then*

$$P(N < \infty) = 1 \quad or \quad P(N = \infty) = 1$$

according as G is or is not a product measure.

REMARK 1. It is known that the bivariate normal distribution with correlation $\rho \ne 1$ is in the domain of attraction of a bivariate extreme-value distribution which is a product measure [Sibuya (1960)]. Thus we get the surprising result that unless $\rho = 1$, the bivariate normal yields a finite number of records.

REMARK 2. Associated with G is a planar Poisson process $\mathscr{P}_G$ that has points at $\{\mathbf{j}_k\}_{k \in \mathbf{N}}$, say, such that $G(\mathbf{x}) = P(\max_k \mathbf{j}_k \le \mathbf{x})$. As above, we may assume the marginal d.f.'s of G are each $\exp(-x^{-1})$ for $x > 0$; the measure ν featured above is then the mean measure of $\mathscr{P}_G$. The probability that $\mathscr{P}_G$ has a greatest point is

$$P(\exists\, k \text{ with } \mathbf{j}_k > \mathbf{j}_l\, \forall\, l \ne k) = \int_{\mathbf{R}^2} e^{-\nu(-\boldsymbol{\infty}, \mathbf{x})^c} \nu(d\mathbf{x}).$$

This is 0 iff ν places no mass in the interior of the support rectangle [cf. Resnick (1987), page 260], which is equivalent to G being a product measure. Thus $P(N = \infty) = 1$ or 0 according as there is positive or zero probability of $\mathscr{P}_G$ having a greatest point.

We thank L. F. M. de Haan for this remark.

6. Postscript. G. O'Brien suggested to us that the set $\mathscr{R}$ of record values be alternatively defined by

$$X_n \in \mathscr{R} \text{ iff } X_k \not\geq X_n, \qquad k = 1, \dots, n-1. \tag{6.1}$$

This makes sense for any partial order and allows more observations to count as records than does (1.1). It is instructive to observe that our first main result, Theorem 2.1, extends to this definition. Thus we redefine H by

$$dH(x) := \frac{d\mu(x)}{\mu[x, \infty)}.$$

Proposition 1.1 need no longer hold, and H can even be non σ-finite. (Consider a nonatomic probability law in $\mathbb{R}^2$ that is concentrated on the line $x + y = 0$. So every X_n is a record.) But we still have $EN_A = H(A)$, and in the proof of Theorem 2.1 the same idea as was used there will establish the Kochen–Stone condition: For $m > n$, $X_m \in A \cap \mathscr{R}$ implies X_m is a record among $X_{n+1}, \dots, X_m$. Hence Theorem 2.1 holds.

Most of our further results do not extend to definition (6.1) because there is no "spatial Markov property" (Theorems 2.2 and 2.6) and we do not have the property $[y, \infty)^c \cap [x, \infty)^c = [x, \infty)^c$ when $x \in [y, \infty)^c$, as would be needed for formulae (2.5) and (2.7).

Acknowledgment. We thank John Haigh for identifying the Stirling numbers in (2.2), and G. O'Brien, E. Perkins, W. Vervaat and a referee for Proposition 1.1.

REFERENCES

AHMED, A. N., ALZAID, A., BARTOSZEWICZ, J. and KOCHAR, S. C. (1986). Dispersive and superadditive ordering. *Adv. in Appl. Probab.* **18** 1019–1022.

BIRKHOFF, G. (1967). *Lattice Theory. Amer. Math. Soc., Colloq. Publ.* **25**. Amer. Math. Soc., Providence, R.I.

CHOW, Y. S. and TEICHER, H. (1978). *Probability Theory: Independence, Interchangeability, Martingales*. Springer, New York.

CHUNG, K. L. (1974). *A Course in Probability Theory*, 2nd ed. Academic, New York.

DE HAAN, L. and RESNICK, S. I. (1977). Limit theory for multivariate sample extremes. *Z. Wahrsch. verw. Gebiete* **40** 317–333.

DEKEN, J. (1976). On records: Scheduled maxima sequences and largest common subsequences. Ph.D. dissertation, Dept. Statistics, Stanford Univ.

DELLACHERIE, C. and MEYER, P.-A. (1978). *Probabilities and Potential.* North-Holland, Amsterdam.

GERRITSE, G. (1985). Lattice-valued semicontinuous functions. Report 8532, Mathematisch Instituut, Katholieke Universiteit Nijmegen.

GIERZ, G., HOFMANN, K. H., KEIMEL, K., LAWSON, J. D., MISLOVE, M. and SCOTT, D. S. (1980). *A Compendium of Continuous Lattices*. Springer, New York.

GILL, R. D. and JOHANSEN, S. (1987). Product integrals and counting processes. Report MS-R8707, CWI, Postbus 4079, 1009AB Amsterdam.

GOLDIE, C. M. and ROGERS, L. C. G. (1984). The k-record processes are i.i.d. *Z. Wahrsch. verw. Gebiete* **67** 197–211.

HALMOS, P. R. (1950). *Measure Theory.* Van Nostrand-Reinhold, New York.

KAMAE, T., KRENGEL, U. and O'BRIEN, G. L. (1977). Stochastic inequalities on partially ordered spaces. *Ann. Probab.* **5** 899–912.

MATHERON, G. (1975). *Random Sets and Integral Geometry*. Wiley, New York.

NACHBIN, L. (1965). *Topology and Order*. Van Nostrand-Reinhold, New York.

NORBERG, T. (1986). Random capacities and their distributions. *Probab. Theory Related Fields* **73** 281–297.

RESNICK, S. I. (1987). *Extreme Values, Regular Variation and Point Processes*. Springer, New York.

RIORDAN, J. (1978). *An Introduction to Combinatorial Analysis*. Princeton Univ. Press, Princeton, N.J.

SHANBHAG, D. N. (1979). Some refinements in distribution theory. *Sankhyā Ser. A* **41** 252–262.

SHORROCK, R. W. (1970). Caravans in traffic flow. Technical Report 20, Dept. Statistics, Stanford Univ.

SHORROCK, R. W. (1972). On record values and record times. *J. Appl. Probab.* **9** 316–326.

SHORROCK, R. W. (1974). On discrete-time extremal processes. *Adv. in Appl. Probab.* **6** 580–592.

SIBUYA, M. (1960). Bivariate extreme statistics. *Ann. Inst. Statist. Math.* **11** 195–210.

STOYAN, H. and STOYAN, D. (1980). On some partial orderings of random closed sets. *Math. Operationsforsch. Statist. Ser. Optim.* **11** 145–154.

VERVAAT, W. (1986). Stationary self-similar extremal processes and random semicontinuous functions. In *Dependence in Probability and Statistics* (E. Eberlein and M. Taqqu, eds.) 457–473. Birkhäuser, Boston.

VERVAAT, W. (1988a). Narrow and vague convergence of set functions. *Statist. Probab. Lett.* **6** 295–298.

VERVAAT, W. (1988b). Random upper semicontinuous functions and extremal processes. Report MS-R8801, CWI, Postbus 4079, 1009AB Amsterdam.

ZIJLSTRA, M. and DE KROON, J. P. M. (1981). Stochastic orderings related to ranking an nth-order record process. *J. Statist. Plann. Inference* **5** 79–81.

STATISTICAL LABORATORY
CAMBRIDGE UNIVERSITY
16 MILL LANE
CAMBRIDGE CB2 1SB
ENGLAND

DEPARTMENT OF OPERATIONS RESEARCH
UPSON HALL
CORNELL UNIVERSITY
ITHACA, NEW YORK 14853-7501