Structure of Record Observations

Charles M. Goldie

Mathematics Department
University of Sussex

IPM, Tehran, 27 October 2011

Contents I

Abstract I

Abstract

This talk is based on the first half of [Bunge & Goldie, 2001], plus some more recent material. Its view of the subject, Records, is almost disjoint from that of other treatments such as [Arnold, Balakrishnan & Nagaraja, 1998] or [Nevzorov, 2000].

1. Record values: Basics I

Notation & basic setup

- F is the distribution function (d.f.) of a random variable (r.v.) X on $\mathbb{R}$.
- $F(x) := P(X \leq x)$, right-continuous.
- $x_+ := \sup\{x : F(x) < 1\}$.
- $\bar{F}(x) := 1 - F(x)$.
- The probability space is assumed to support X, X_1, X_2, $\ldots$, which are i.i.d. (independent and identically distributed) $\sim F$.
- The *order statistics* $X_n^1 \geq X_n^2 \geq \cdots \geq X_n^n$ are $X_1, \ldots, X_n$ in order.

1. Record values: Basics I

Definition 1.1.

The *initial rank* of X_n is $\rho_n := \sum_{i=1}^{n} \mathbf{1}_{X_i \geq X_n}$.

Definition 1.2.

X_n is a *k-record* if $\rho_n = k$.
The *record values* are the 1-records.

- X_1 is a record value.
- For $k > 1$ the first k-record occurs at or after time k.
- These are *upper* records: one can alternatively work with lower records.

Definition 1.3.

Let $R_1^k < R_2^k < \cdots$ denote the successive k-records (note the strict inequalities). The whole sequence is denoted $\mathbf{R}^k = (R_1^k, R_2^k, \dots)$.

1. Record values: Structure I

We determine the structure of $\mathbf{R}^1$. Later this will yield the structure of $\mathbf{R}^k$ for every k.

If $x_+ < \infty$ and $P(X = x_+) > 0$ there will be a final k-record; otherwise not.

Because the (finite or infinite) sequence $\mathbf{R}^1$ is strictly increasing we may regard it as a (random) set.

Lemma 1.4.

Let E be any finite union of (disjoint) intervals $(u, v]$ in $(-\infty, x_+]$. Then

$$P(\mathbf{R}^1 \cap E = \emptyset) = e^{-\eta(E)} \tag{1}$$

where η is the measure on $(-\infty, x_+]$ defined by

$$\eta(-\infty, x] := -\ln \bar{F}(x).$$

If $x_+ < \infty$ and $P(X = x_+) > 0$ then $\eta(-\infty, x_+) < \infty$ and $\eta\{x_+\} = \infty$. Otherwise, and in particular if $x_+ = \infty$, $\eta(-\infty, x_+) = \infty$ and $\eta\{x_+\} = 0$. Always, $\eta(-\infty, x] < \infty$ for all $x < x_+$.

Definition 1.5.

η is the *avoidance measure* of $\mathbf{R}^1$.

1. Record values: Structure II

Theorem 1.6.

The law of $\mathbf{R}^1$ *is the unique law of a simple*[1] *point process such that* (1) *holds for all finite unions* E *of intervals.*

Proof.

From point-process theory. □

Let D be the set of points where F is discontinuous.

D is also the set of atoms of η, i.e. the set of points where $\eta(-\infty, x]$ is discontinuous.

Because $\bar{F}(x) = F(x, \infty) = e^{-\eta(-\infty, x]}$,

$$\therefore \qquad F[x, \infty) = e^{-\eta(-\infty, x)};$$

$$\therefore \qquad \frac{F(x, \infty)}{F[x, \infty)} = e^{-\eta\{x\}} = P(\mathbf{R}^1 \cap \{x\} = \emptyset).$$

Definition 1.7.

The *discrete part* of η is the measure $\eta_d(E) := \sum_{x \in D \cap E} \eta\{x\}$.
The *continuous part* of η is the measure $\eta_c := \eta - \eta_d$.

[1] simple: no multiple points

1. Record values: Structure I

Theorem 1.8 ([Shorrock, 1972], [Shorrock, 1974]).

$\mathbf{R}^1$ is composed of

- *a Poisson process $\mathbf{R}^1_c$ of characteristic measure η_c,*

and, independently of $\mathbf{R}^1_c$ and of each other,

- *at each $x \in D$, a demon who*
  - *with probability $1 - e^{-\eta\{x\}}$ gives $\mathbf{R}^1$ a point at x,*
  - *or with probability $e^{-\eta\{x\}}$ does not.*

Also $\mathbf{R}^1$ is completely random (= independent increments) and satisfies (1) *for all Borel sets E.*

Note

A *Poisson process* as referred to above is more precisely an *inhomogeneous Poisson process of continuous characteristic measure* $\nu = \eta_c$.

This is a simple, completely random point process N with, for any Borel set B, $N(B) \sim \mathrm{Pois}(\nu(B))$.

1. Record values: Ignatov's Theorem I

Theorem 1.9 ('Ignatov').
$\mathbf{R}^1, \mathbf{R}^2, \ldots$ *are i.i.d.*

Proof history

- [Ignatov, 1976/77], submitted 1978, appeared 1986; continuous case
- [Deheuvels, 1983], continuous case; incomplete
- [Goldie, 1983]
- [Goldie & Rogers, 1984]
- [Stam, 1985], continuous case
- [Engelen, Tommassen & Vervaat, 1988]
- [Samuels, 1992]
- [Yao, 1997]
- [Gnedin, 2008], continuous case

2. Record times: Sojourns I

Assume $P(X = x_+) = 0$, so each $\mathbf{R}^k$ is an infinite sequence.
Fix $k \in \mathbb{N}$.
Arrange $\mathbf{Q}^k := \mathbf{R}^1 \cup \cdots \cup \mathbf{R}^k$ in a sequence in increasing order:

$$\mathbf{Q}^k = \{Q_1^k \leq Q_2^k \leq \cdots\}.$$

If F is not continuous this sequence can contain repeats.
Let $L_1^k := k$ and

$$L_{j+1}^k := \min\{n : n > L_j^k,\ \rho_n \leq k\}.$$

These are the times when $X_\cdot^k$ steps to the next point of $\mathbf{Q}^k$:

$$X_n^k = Q_j^k \text{ for all } n \text{ with } L_j^k \leq n < L_{j+1}^k.$$

Definition 2.1.
The *sojourn* of the k^{th} order statistic at Q_j^k, the time it spends there, is

$$\Delta_j^k := L_{j+1}^k - L_j^k \qquad (j = 1, 2, \ldots).$$

2. Record times: Sojourns I

Theorem 2.2.

The sojourns Δ_1^k, Δ_2^k, ... are conditionally independent given $\mathbf{Q}^k$, with geometric distributions

$$P(\Delta_j^k = l|\mathbf{Q}^k) = \left(F(Q_j^k)\right)^{l-1} \bar{F}(Q_j^k) \qquad (l = 1, 2, \ldots).$$

Definition 2.3.

Let $X^{m\leftarrow}(\cdot)$ be the left-continuous inverse of $X_\cdot^m$:

$$X^{m\leftarrow}(x) := \inf\{n \geq m : X_n^m \geq x\} \qquad (x \leq x_+).$$

Left-continuity yields the convenient relationship

$$X^{m\leftarrow}(x) \leq n \quad \text{iff} \quad x \leq X_n^m.$$

Then

$$\begin{aligned} X^{m\leftarrow}[x, y) &:= X^{m\leftarrow}(y) - X^{m\leftarrow}(x) \\ &= \#\{n : X_n^m \in [x, y)\}. \end{aligned}$$

2. Record times: Sojourns II

Theorem 2.4.

Fix $m \in \mathbb{N}$. The process $X^{m\leftarrow}$ has independent increments: for any disjoint intervals $I_1, \ldots, I_k$ in $(-\infty, x_+]$,

$$P(X^{m\leftarrow} I_1 = n_1, \ldots, X^{m\leftarrow} I_k = n_k) = \prod_{l=1}^{k} P(X^{m\leftarrow} I_l = n_l),$$

and

$$P(X^{m\leftarrow}[x,y) = n) = \begin{cases} \left(\frac{F[y,\infty)}{F[x,\infty)}\right)^m & \text{for } n = 0, \\ \left(\frac{F[y,\infty)}{F[x,\infty)}\right)^m \sum_{k=1}^{m\wedge n} \binom{m}{k}\binom{n-1}{k-1} F[x,y)^k F(-\infty,y)^{n-k} & \\ \quad \text{for } n = 1, 2, \ldots. \end{cases} \tag{2}$$

2. Record times: Times I

Theorem 2.5 (Dwass-Rényi Lemma: [Dwass, 1960], [Rényi, 1962]).

Assume F continuous. Then ρ_1, ρ_2, ... are independent and $\rho_n \sim \text{Unif}\{1, \ldots, n\}$.

The uniform distribution here is *discrete uniform*.

Proof.
Exercise! □

Aside: number of records

Let N_n be the number of records among X_1, ..., X_n:

$$N_n := \sum_{k=1}^{n} I_k \quad \text{where} \quad I_k := \mathbf{1}\{\rho_k = 1\}.$$

By the Dwass-Rényi Lemma the I_k are *independent* with $P(I_k = 1) = k^{-1}$, $P(I_k = 0) = 1 - k^{-1}$.

2. Record times: Times I

So, with $\gamma = 0{\cdot}57721\cdots$ the Euler-Mascheroni constant,

$$EN_n = \sum_1^k \frac{1}{k} = \ln n + \gamma + \mathrm{O}\left(\frac{1}{n}\right);$$

$$\operatorname{var} N_n = \sum_1^n \left(\frac{1}{k} - \frac{1}{k^2}\right) = \ln n + \gamma - \frac{\pi^2}{6} + \mathrm{O}\left(\frac{1}{n}\right).$$

One may prove

$$\frac{N_n}{\ln n} \xrightarrow{a.s.} 1, \quad \frac{N_n - \ln n}{\sqrt{\ln n}} \Longrightarrow \mathrm{N}(0,1),$$

etc.

2. Record times: Times I

Definition 2.6.

The values of n when $\rho_n = 1$ are the *record times* $1 = L_1 < L_2 < \cdots$.

If $P(X = x_+) = 0$ this is an infinite sequence.

We restrict attention to record times but all that follows has versions for k-record times, suitably defined.

Theorem 2.7.

Assume F continuous. Then $(L_n)_{n\geq 1}$ is a Markov chain with $L_1 = 1$ and stationary transition laws

$$
\begin{aligned}
P(L_{n+1} = l | L_n = j) &= \frac{j}{(l-1)l} \quad (l = j+1, j+2, \dots) \\
&= \frac{j}{l-1} - \frac{j}{l}.
\end{aligned} \tag{3}
$$

Lemma 2.8.

Assume F continuous. Let W_1, W_2, ... be i.i.d. Unif$(0,1)$, *independent of $(L_j)_{j\geq 1}$. Define*

$$Y_n := -\ln\left((1-W_n)\frac{L_n}{L_{n+1}} + W_n\frac{L_n}{L_{n+1}-1}\right), \quad (n = 1, 2, \ldots).$$

Then Y_1, Y_2, ... are i.i.d. Expon(1) *r.v.s.*

Theorem 2.9 (Williams-Pfeifer Strong Approximation for Record Times, [Williams, 1973], [Pfeifer, 1987]).

Assume F continuous. Use the probability space extended by the W_n as above. Then

$$L_{n+1} = \lceil L_n e^{Y_n} \rceil \quad \textit{for } n = 1, 2, \ldots,$$

where

$$\lceil x \rceil := \min\{n \textit{ integer}, \; n \geq x\}.$$

Theorem 2.10 ([Pfeifer, 1987]).

Assume F continuous. Use the probability space extended by the W_n as in Lemma 2.8. Set $S_n := \sum_1^n Y_j$. Then there exists $Z > 0$ with $E(Z^k) < \infty$ for all k, such that Z and $(S_n - n)/\sqrt{n}$ are asymptotically independent, and

$$\ln L_n = Z + S_{n-1} + o(e^{-n/2}) \ a.s. \quad (n \to \infty).$$

3. Limits: Limit laws for record values I

Notation

$\xrightarrow{L}$ denotes convergence in law (in distribution), and $\overset{L}{=}$ denotes equality of probability laws (distributions).

Definition 3.1.

R.v.s X, Y, or equivalently their laws F, G, are *of the same type* if there exist $a \in (0, \infty)$ and $b \in \mathbb{R}$ so that

$$Y \overset{L}{=} aX + b, \text{ equivalently } G(y) = F\left(\frac{y-b}{a}\right) \; \forall y.$$

This is an equivalence relation on laws on $\mathbb{R}$ (exercise). The equivalence classes are the *types*.

3. Limits: Limit laws for record values I

Theorem 3.2 (Convergence of Types).

Let X, X_n be r.v.s, $a_n > 0$, $b_n \in \mathbb{R}$ (norming and centring constants, or scale and location constants), such that

$$\frac{X_n - b_n}{a_n} \xrightarrow{L} X \qquad (n \to \infty),$$

with X non-degenerate. Let Y be a r.v., $\alpha_n > 0$, $\beta_n \in \mathbb{R}$ constants. Then

$$(i)\ \frac{X_n - \beta_n}{\alpha_n} \xrightarrow{L} Y$$

iff

$$(ii)\ \frac{a_n}{\alpha_n} \to \alpha \in [0, \infty),\ \frac{b_n - \beta_n}{\alpha_n} \to \beta \in \mathbb{R} \quad (n \to \infty).$$

In that case $Y \stackrel{L}{=} \alpha X + \beta$, and $\alpha \geq 0$, β are the unique constants for which this holds.

When (i) or (ii) holds, Y is non-degenerate iff $\alpha > 0$, and X and Y are then of the same type.

3. Limits: Limit laws for record values I

Theorem 3.3 ([Resnick, 1973]).

Assume F continuous. Then the possible limit laws for $(R_n - b_n)/a_n$ are those in the type of one of

(i) $\tilde{\Phi}_\alpha(x) := \begin{cases} 0 & \text{if } x \le 0, \\ \Phi(\ln x^\alpha) & \text{if } x > 0; \end{cases}$

(ii) $\tilde{\Psi}_\alpha(x) := \begin{cases} \Phi\big(\ln(-x^\alpha)\big) & \text{if } x < 0, \\ 1 & \text{if } x \ge 0; \end{cases}$

(iii) Φ,

where Φ is the $\mathrm{N}(0,1)$ *d.f. and $\alpha > 0$ is constant.*

Resnick also characterises of the *domain of attraction for record values* for each of these limit types. That is, for G each of $\tilde{\Phi}_\alpha$, $\tilde{\Psi}_\alpha$, Φ, he finds those F for which there exist $a_n > 0$ and b_n such that $(R_n^1 - b_n)/a_n \xrightarrow{L} G$.

4. Extensions I

Definition 4.1.

$H(x) := H(-\infty, x]$ where H is the *hazard measure* corresponding to F, defined by

$$H(A) := \int_A \frac{dF(x)}{F[x, \infty)}$$

for Borel sets A in $\mathbb{R}$.

Proposition 4.2.

H is the intensity measure of the point process $\mathbf{R}^1$: $H(A) = E\#(\mathbf{R}^1 \cap A)$.

Proposition 4.3.

Assume F continuous. Then

$$\big(H(R_n)\big)_{n \geq 1} \stackrel{L}{=} \left(\sum_1^n E_i\right)_{n \geq 1},$$

where $E_1, E_2, \ldots$ are independent Expon(1) *r.v.s.*

These results suggest how one might generalise records to multidimensional and other general settings.

4. Extensions: Records in a poset I

Let S be a set with a partial order $<$.
That is, the *graph* $G_< := \{(x, y) \in S \times S : x < y\}$ has properties

(i) antisymmetry: $x < x$ for *no* $x \in S$,

(ii) transitivity: $x < y,\ y < z \xrightarrow{L} x < z$.

Let $\mathcal{S}$ be a σ-algebra of subsets of S. Assume

$$G_< \in \mathcal{S} \times \mathcal{S} \quad \text{(the graph is product-measurable).}$$

Let μ be a probability law on $(S, \mathcal{S})$.
Let $X_1, X_2, \ldots$ be i.i.d. $\sim \mu$.
Adjoin extra points $-\infty$, ∞ with the properties

$$-\infty < x < \infty \quad \forall\, x \in S.$$

Let $S^* := S \cup \{-\infty, \infty\}$.
Define *intervals* $(x, y) := \{z \in S : x < z < y\}$ for $x, y \in S^*$.
Let

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

4. Extensions: Records in a poset I

Proposition 4.4.
$\mu(S_\mu) = 1$.

Proof.
Fubini! □

So S_μ functions as the 'support' of μ.

Definition 4.5.
Hazard measure H on $(S, \mathcal{S})$:

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

Definition 4.6.
X_n is a *record* if $X_k < X_n$ for $k = 1, \dots, n-1$.

Let R denote the set of records.

4. Extensions: Records in a poset I

Theorem 4.7 ([Goldie & Resnick, 1989]).

Let $A \in \mathcal{S}$ and define events $A_n := \{X_n \in A \cap R\}$. Then $\sum_{n=1}^{\infty} P(A_n) = H(A)$. Further, $P(\#(R \cap A) = \infty) = 1$ or 0 according as $\sum_{n=1}^{\infty} P(A_n) = H(A) = \infty$ or $< \infty$.

4. Extensions: Strict multivariate records I

Continue with the setup and notation of the last section, but now specialise to $\mathbb{R}^d$ with $d > 1$, define $x < y$ component-wise, and take $A := \mathbb{R}^d$.

Theorem 4.8 ([Gnedin, 1998]).

If F is a non-singular Gaussian law on $\mathbb{R}^d$, with correlation matrix Λ, then there exist $\alpha > 1$ and $\beta \in \{2, \ldots, d\}$, both depending on Λ, so that

$$P(A_n) \asymp n^{-\alpha} (\ln n)^{(\alpha-\beta)/2}.$$

Consequently $P(\#R < \infty) = 1$ for all non-singular Gaussian laws. The same holds for singular Gaussian laws unless all correlation coefficients are $+1$.

For $d = 2$ and correlation coefficient $\rho \in (-1, 1)$, more precisely,

$$P(A_n) \asymp n^{-2/(1+\rho)} (\ln n)^{-\rho/(1+\rho)}.$$

4. Extensions: Chain records I

Let $X, X_1, X_2, \ldots$ be i.i.d. in $\mathbb{R}^d$. Define $x < y$ component-wise. Define a form of *lower* record as follows:

- X_n is a *chain record* if it is below the previous chain record.

Definition 4.9.
Set $T_1 := 1$, and for $k = 2, 3, \ldots,$

$$T_k := \min\{n > T_{k-1} : X_n < X_{T_{k-1}}\}.$$

The *chain records* are $R_k := X_{T_k}$.

Let $N_n := \sum_{j=1}^{n} \mathbf{1}\{X_j \text{ is a chain record}\}$.

4. Extensions: Chain records I

Theorem 4.10 ([Gnedin, 2007]).

Suppose X has a continuous product distribution. Let W be the product of d independent $\mathrm{Unif}(0,1)$ *r.v.s, so that* $m := E(-\ln W) = d$ *and* $\sigma^2 := \mathrm{var}(-\ln W) = d$*. Then* $N_n \sim m^{-1} \ln n$ *a.s. and*

$$\frac{N_n - m^{-1} \ln n}{\sqrt{\sigma^2 m^{-3} \ln n}} \xrightarrow{L} \mathrm{N}(0,1).$$

Note

The $d = 1$ case is included! For $d = 1$, $W \sim \mathrm{Unif}(0,1)$ so $-\ln W \sim \mathrm{Expon}(1)$, so $m = 1 = \sigma^2$. As in §2,

$$N_n \sim \ln n \ a.s., \qquad \frac{N_n - \ln n}{\sqrt{\ln n}} \xrightarrow{L} \mathrm{N}(0,1).$$

4. Extensions: Longest sequences I

In $\mathbb{R}^2$ use the notation $\mathbf{x} = (x_1, x_2)$, define $\mathbf{x} < \mathbf{y}$ component-wise, take $\mathbf{X}_1, \mathbf{X}_2, \ldots$ i.i.d. $\sim F$, and as above define hazard measure H by

$$H(d\mathbf{x}) := \frac{F(d\mathbf{x})}{1 - F(\mathbf{x}-)} = \frac{P(\mathbf{X}_1 \in d\mathbf{x})}{P(\{\mathbf{X}_1 < \mathbf{x}\}^c)}.$$

Let A be an interval $[\mathbf{a}, \mathbf{b}]$ in $\mathbb{R}^2$. Considering strict records (in both coordinates simultaneously) we know from Theorem 4.7 that the number N_A of records falling in A is finite a.s. if and only if $H(A) < \infty$. In this section we will find out about the r.v. N_A, when it is finite.

Given points $\mathbf{x}_1 < \cdots < \mathbf{x}_n$ in A, join $\mathbf{a} < \mathbf{x}_1 < \cdots < \mathbf{x}_n < \mathbf{b}$ by straight lines to form a path.

4. Extensions: Longest sequences I

Theorem 4.11 ([Goldie & Resnick, 1995]).

Assume $H(A) < \infty$, that H has a bounded density on A and that the distribution G on A given by

$$G(\mathbf{x}) := \frac{H(\mathbf{x})}{H(A)} \qquad (\mathbf{x} \in A)$$

satisfies the conditions of either Theorem 4.13 or Theorem 4.21 below. Then, given $N_A = n$, as $n \to \infty$ the path joining the records converges in probability to a non-random limit curve which maximises the Deuschel-Zeitouni functional $J(\phi)$ or the Goldie-Resnick functional $T(\mathbf{f})$ respectively.

4. Extensions: Longest sequences I

Definition 4.12.

Let $B^{\uparrow}$ be the set of non-decreasing right-continuous functions $\phi : [a_1, a_2] \to [b_1, b_2]$. For $\phi \in B^{\uparrow}$, $\phi(x) = \int_0^x \dot{\phi}(t)\,dt + \phi_s(x)$ where ϕ_s is singular. Assuming G has a density g, define $J : B^{\uparrow} \to \mathbb{R}$ by

$$J(\phi) := \int_{a_1}^{a_2} \sqrt{\dot{\phi}(x) g\big(x, \phi(x)\big)}\, dx.$$

Theorem 4.13 ([Deuschel & Zeitouni, 1995]).

Let $\mathbf{Z}_1, \dots, \mathbf{Z}_n$ be i.i.d. $\sim G$ on an interval A in $\mathbb{R}^2$. Assume that G has a density g that is C_b^1 and such that $\ln g$ is bounded. Assume also that

$$J(\phi) \textit{ is maximised on a finite set } \{\bar{\phi}_1, \dots, \bar{\phi}_k\}. \tag{4}$$

On the event $\mathbf{Z}_1 < \cdots < \mathbf{Z}_n$ let ϕ_n denote the element of $B^{\uparrow}$ formed by joining $\mathbf{a} < \mathbf{Z}_1 < \cdots < \mathbf{Z}_n < \mathbf{b}$ by straight-line segments. Then for each $\varepsilon > 0$,

$$P(\min\{\|\phi_n - \bar{\phi}_1\|_\infty, \dots, \|\phi_n - \bar{\phi}_k\|_\infty\} > \varepsilon | \mathbf{Z}_1 < \cdots < \mathbf{Z}_n) \to 0 \quad (n \to \infty).$$

4. Extensions: Longest sequences II

Set $\bar{J} := \sup_{\phi \in B\uparrow} J(\phi)$. When G is a product distribution the diagonal is the unique maximising curve $\bar{\phi}$, and then obviously $\bar{J} = J(\bar{\phi}) = 1$.

Definition 4.14.

For $y_1, \ldots, y_n \in \mathbb{R}$ an *increasing subsequence* is $y_{i_1} < y_{i_2} < \cdots < y_{i_k}$ where $i_1 < i_2 < \cdots < i_k$.

(That is, in selecting the y you can miss indices out: the y selected don't have to be a *run*.)

Theorem 4.15 ([Deuschel & Zeitouni, 1995]).

Let $\mathbf{Z}_1, \ldots, \mathbf{Z}_n$ be i.i.d. $\sim G$ in $[0,1]^2$. Order them by their x components and let L_n be the length of the longest increasing subsequence (of the y components). Assume that G has a density g that is C^1_b and such that $\ln g$ is bounded. Then $L_n/\sqrt{n} \xrightarrow{P} 2\bar{J}$.

4. Extensions: Longest sequences I

This builds on, and extends, the celebrated solution to *Ulam's problem*:

Theorem 4.16 ([Vershik & Kerov, 1977], [Logan & Shepp, 1977], [Aldous & Diaconis, 1995], [Seppäläinen, 1996],).

Let l_n be the length of the longest increasing subsequence in a random permutation of order n. Then $l_n/\sqrt{n} \xrightarrow{P} 2$.

Proof.

See the cited references, or for a survey [Aldous & Diaconis, 1999]. $\square$

4. Extensions: Longest sequences I

To avoid the rather unsatisfactory condition (4), different assumptions seem to be needed, and will lead to further conclusions. First, an important concept from information theory.

Definition 4.17.

For probability measures μ, ν on a common measurable space the *I-divergence* (*Kullback-Leibler information number*, *relative entropy*) is

$$D(\mu\|\nu) := \begin{cases} \int \left(\ln \frac{d\mu}{d\nu}\right) d\mu & \text{if } \mu \ll \nu, \\ \infty & \text{if not.} \end{cases}$$

For probability densities p, q on $\mathbb{R}$ this reduces to

$$D(p\|q) := \begin{cases} \int_{-\infty}^{\infty} \left(\ln \frac{p(x)}{q(x)}\right) p(x)\, dx & \text{if } p(x) = 0 \text{ whenever } q(x) = 0, \\ \infty & \text{if not.} \end{cases}$$

Fact 4.18.

$0 \le D(\mu\|\nu) \le \infty$.

4. Extensions: Longest sequences I

Aside: Statistics

Theorem 4.19 (Stein's Lemma or the 'Chernoff-Stein Lemma').

For testing

$$H_0 : \text{the density is } p,$$

against

$$H_1 : \text{the density is } q,$$

the most powerful level-α test, based on a random sample of size n, has Type II error probability

$$\beta_n(\alpha) = e^{-D(p\|q)n(1+o(1))} \text{ as } n \to \infty.$$

Proof.
See e.g. [Cover & Thomas, 2006]. □

4. Extensions: Longest sequences II

Definition 4.20.

A *parametrised curve* $\mathbf{f} = (f_1, f_2)$ on $A = [\mathbf{a}, \mathbf{b}] \subset \mathbb{R}^2$ is an element of D_L, the space of non-decreasing functions $\mathbf{f} : [0,1] \to \mathbb{R}^2$ that are left-continuous on $(0,1]$ and have $\mathbf{f}(0) = \mathbf{a}$ and $\mathbf{f}(1) \leq \mathbf{b}$. Assuming G has a density g, define $T : D_L \to \mathbb{R}$ by

$$T(\mathbf{f}) := \int_0^1 \ln g\big(\mathbf{f}(p)\big)\, dp - D\Big(f_1(U) \Big\| U\Big) - D\Big(f_2(V) \Big\| V\Big) \tag{5}$$

where $U \sim \mathrm{Unif}(a_1, a_2)$ and $V \sim \mathrm{Unif}(b_1, b_2)$.

4. Extensions: Longest sequences I

Theorem 4.21 ([Goldie & Resnick, 1995]).

Assume that G has a density g that is continuous and has $\ln g$ bounded on A, and further is such that $\ln g$ is L-superadditive on A°:

$$\frac{\partial^2 \ln g(x,y)}{\partial x \partial y} \geq 0 \quad ((x,y) \in A^\circ) \tag{6}$$

(g is 'humped'). Then there is a unique $\bar{\mathbf{f}} \in D_L$ that maximises $T(\mathbf{f})$.

Let $\mathbf{Z}_1, \ldots, \mathbf{Z}_n$ be i.i.d. $\sim G$ on A. On the event $\mathbf{Z}_1 < \cdots < \mathbf{Z}_n$ let $\mathbf{f}_n$ denote the element of D_L formed by joining $\mathbf{a} < \mathbf{Z}_1 < \cdots < \mathbf{Z}_n < \mathbf{b}$ by straight-line segments. Then

$$\|\mathbf{f}_n - \bar{\mathbf{f}}\|_\infty \xrightarrow{P} 0 \quad (n \to \infty)$$

(where $\|\mathbf{f}\|_\infty := \sup_{p \in [0,1]} \|\mathbf{f}(p)\|$ for $\mathbf{f} : [0,1] \to \mathbb{R}^2$, and $\|\cdot\|$ is any norm on $\mathbb{R}^2$).

4. Extensions: Longest sequences I

Fact 4.22.
$T(\bar{\mathbf{f}}) = 2 \ln \bar{J}$.

Theorem 4.23 ([Goldie & Resnick, 1995]).
Under the conditions of Theorem 4.21,

$$P(\mathbf{Z}_1 < \cdots < \mathbf{Z}_n) = e^{-n(2\ln n - 2 - \ln|A| - T(\bar{\mathbf{f}}) + o(1))} \quad (n \to \infty)$$

and

$$P(\mathbf{Z}_1, \ldots, \mathbf{Z}_n \text{ can be ordered}) = e^{-n(\ln n - 1 - \ln|A| - T(\bar{\mathbf{f}}) + o(1))} \quad (n \to \infty),$$

where $|A| = \mathrm{Leb}(A) = (b_1 - a_1)(b_2 - a_2)$.

4. Extensions: Longest sequences I

Theorem 4.24 ([Goldie & Resnick, 1995]).

Let $A = [\mathbf{a}, \mathbf{b}]$ *be an interval in* $\mathbb{R}^2$*. Let* $\mathbf{X}_1, \mathbf{X}_2, \ldots$ *be i.i.d.* $\sim F$ *where* F *has continuous density* f*. Set* $h(\mathbf{x}) := f(\mathbf{x})/\big(1 - F(\mathbf{x})\big)$*,* $H(A) := \int_A h(\mathbf{x})\, d\mathbf{x}$*,* $g(\mathbf{x}) := h(\mathbf{x})/H(A)$ *for* $\mathbf{x} \in A$*, and hence define* $T(\cdot)$ *by* (5)*. Suppose further that* $F(\mathbf{b}) < 1$*,* $\ln f$ *is bounded in* A*, and that* $\ln h$ *is L-superadditive (see* (6)*) on* A°*. Then*

$$P(N_A = n) = \frac{\left(|A|e^{T(\bar{\mathbf{f}})} + \mathrm{o}(1)\right)^n}{(n!)^2} \quad (n \to \infty)$$

and

$$P(N_A \geq n) = \frac{\left(|A|e^{T(\bar{\mathbf{f}})} + \mathrm{o}(1)\right)^n}{(n!)^2} \quad (n \to \infty).$$

References I

[Aldous & Diaconis, 1995] D. J. Aldous and P. Diaconis, 'Hammersley's interacting particle process and longest increasing sequences'. *Probab. Th. Rel. Fields* **103**, 199–213.

[Aldous & Diaconis, 1999] D. J. Aldous and P. Diaconis, 'Longest increasing subsequences: from patience sorting to the Baik-Deift-Johansson theorem'. *Bull. Amer. Math. Soc. (N.S.)* **36**, 413–432.

[Arnold, Balakrishnan & Nagaraja, 1998] B. C. Arnold, N. Balakrishnan and H. N. Nagaraja, *Records*. Wiley-Interscience, New York. ISBN13 978–0471081081.

[Bunge & Goldie, 2001] J. Bunge and C. M. Goldie, 'Record sequences and their applications'. In *Handbook of Statistics* vol. 19 (Stochastic Processes, Theory and Methods), ed. D. N. Shanbhag, C. R. Rao, pp. 277–308. North-Holland, Amsterdam. ISBN10 0–444–50014–6.

[Cover & Thomas, 2006] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Wiley, New York. ISBN13 9780471241959.

References II

[Deheuvels, 1983] P. Deheuvels, 'The strong approximation of extremal processes, II'. *Z. Wahrscheinlichkeitstheorie u. verw. Gebiete* **62**, 7–15.

[Deuschel & Zeitouni, 1995] J.-D. Deuschel and O. Zeitouni, 'Limiting curves for i.i.d. records'. *Ann. Probab.* **23**, 852–878.

[Dwass, 1960] M. Dwass, 'Some k-sample rank order tests'. In *Contributions to Probability and Statistics: Essays in Honor of H. Hotelling*, ed. I. Olkin, S. G. Ghurye, W. Hoeffding, W. G. Madow, H. B. Mann, pp. 198–202. Stanford Studies in Math. & Stat. **2**, Stanford U. Press, Stanford.

[Engelen, Tommassen & Vervaat, 1988] R. Engelen, P. Tommassen and W. Vervaat, 'Ignatov's theorem: a new and short proof'. *J. Appl. Probab.* **25A**, 229–236.

[Gnedin, 1998] A. V. Gnedin, 'Records from a multivariate normal sample'. *Statist. Probab. Letters*, **39**, 11–15.

[Gnedin, 2007] A. V. Gnedin, 'The chain records'. *Electron. J. Probab.* **12**, 767–786.

References III

[Gnedin, 2008] A. V. Gnedin, 'Corners and records of the Poisson process in quadrant'. *Electron. Comm. Probab.* **13**, 187–193.

[Goldie, 1983] C. M. Goldie, *On Records and Related Topics in Probability Theory*. Report, School of Mathematical & Physical Sciences, University of Sussex, UK.

[Goldie & Resnick, 1989] C. M. Goldie and S. I. Resnick, 'Records in a partially ordered set'. *Ann. Probab.* **17**, 678–689.

[Goldie & Resnick, 1995] C. M. Goldie and S. I. Resnick, 'Many multivariate records'. *Stoch. Proc. Appl.* **59**, 185–216.

[Goldie & Rogers, 1984] C. M. Goldie and L. C. G. Rogers, 'The k-records are i.i.d.' *Z. Wahrscheinlichkeitstheorie u. verw. Gebiete* **67**, 197–211.

[Ignatov, 1976/77] Z. Ignatov, 'Ein von der Variationsreihe erzeugter Poissonscher Punktprozess'. *Annuaire Univ. Sofia Fac. Math. Méc.* **71**, no. 2, 79–94.

[Logan & Shepp, 1977] B. F. Logan and L. A. Shepp, 'A variational problem for random Young tableaux'. *Adv. in Math.* **26**, 206-222.

References IV

[Nevzorov, 2000] V. B. Nevzorov, *Records: Mathematical Theory*. Transl. of Math. Monographs **194**, Amer. Math. Soc., Providence. ISBN10 0–8218–1945–3.

[Rényi, 1962] A. Rényi, ‘On the extreme elements of observations’. *MTA III. Oszt. Közl.* **12**, 105–121. Reprinted in *Selected Papers*, vol. 3, ed. P. Turán. Akadémiai Kiadó, Budapest (1976), 50–65.

[Pfeifer, 1987] D. Pfeifer, ‘On a joint strong approximation theorem for record and inter-record times’. *Probab. Th. Rel. Fields* **75**, 213–221.

[Resnick, 1973] S. I. Resnick, ‘Limit laws for record values’. *Stoch. Proc. Appl.* **1**, 67–82.

[Samuels, 1992] S. M. Samuels, ‘An all-at-once proof of Ignatov’s theorem’. In *Strategies for Sequential Search and Selection in Real Time*, ed. F. T. Bruss, T. S. Ferguson, S. M. Samuels, pp. 231–237. Contemporary Mathematics **125**, Amer. Math. Soc., Providence. ISBN13 978-0-8218-5133-3.

[Seppäläinen, 1996] T. Seppäläinen, ‘A microscopic model for the Burgers equation and longest increasing subsequences’. *Electron. J. Probab.* **1**, 1–51.

References V

[Shorrock, 1972] R. W. Shorrock, 'On record values and record times'. *J. Appl. Probab.* **9**, 316–326.

[Shorrock, 1974] R. W. Shorrock, 'On discrete-time extremal processes'. *Adv. Appl. Probab.* **6**, 580–592.

[Stam, 1985] A. J. Stam, 'Independent Poisson processes generated by record values and inter-record times'. *Stoch. Proc. Appl.* **19**, 315–325.

[Vershik & Kerov, 1977] A. M. Vershik and S. V. Kerov, 'Asymptotics of the Plancherel measure of the symmetric group and the limiting form of Young tables'. *Soviet Math. Dokl.* **18**, 527–531; transl. of *Dokl. Acad. Nauk. SSSR* **233**(1977), 1024–1027.

[Williams, 1973] D. Williams, 'On Rényi's 'record' problem and Engel's series'. *Bull. London Math. Soc.* **5**, 235–237.

[Yao, 1997] Y.-C. Yao, 'On independence of k-record processes: Ignatov's theorem revisited. *Ann. Appl. Probab.* **7**, 815–821.