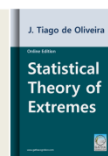




Statistical Theory of Extremes

Homepage: <http://www.gathacognition.com/book/gcb14>
<http://dx.doi.org/10.21523/gcb1>



Part 4

Stochastic sequences and Processes of Extremes

Chapter 14

Extremal Sequences and Processes: Basics and Statistics

J. Tiago de Oliveira

Academia das Ciências de Lisboa (Lisbon Academy of Sciences), Lisbon, Portugal.

Abstract

This chapter reveals a definition of extremal (maxima) random sequences and stochastic processes. Extremal sequences are the usual sequences of maxima. It can be obtained as stochastic processes of independent extremes with extremal margins. Extremal sequences and processes (distribution and properties), limiting characterizations, second order properties of extremal processes, jumps of an extremal process, logarithmization of time and the ergodic theorem for extremal processes, parameterization of extremal processes and statistics for extremal sequences and processes are discussed with explanatory examples.

Published Online

23 June 2017

Keywords

Extremal sequences;
 Extremal processes;
 Maxima;
 Independent extremes;
 Limiting characterizations;
 Parameterization;
 Ergodic theorem.

Editor(s)

J.C. Tiago de Oliveira

14.1 Introduction

We present a definition of extremal (maxima) random sequences and stochastic processes [the latter being distinct from the limiting constructions given by Dwass (1964) and Lamperti (1964)] following Tiago de Oliveira (1968); the insistence will be more on an *actual* characterization than on a limiting one. Extremal sequences are the usual sequences of maxima, and extremal processes will be obtained as stochastic processes (incontinuous time) of independent extremes (maxima) with extremal (maxima) margins. The limiting approaches will be described in a special section.

Originally published in 'Statistical Analysis of Extremes', 1997, 2016

<http://dx.doi.org/10.21523/gcb1.1723>

© 2017 GATHA COGNITION® All rights reserved.

The definition being the weak one (in law), as seen essentially in the identifiability problem, the properties described are, chiefly, the second order properties for the case studied, with Gumbel margins.

It will be clear that extremal processes are surely continuous to the right and almost surely with discontinuous trajectories.

We will always assume the marginal distributions of extremal sequences and processes to be the Gumbel distribution ($\Lambda(z) = (\exp(-e^{-z}))$, for reduced random variables). The other extremal sequences and processes are easily converted to Gumbel margins by logarithmic transformations if they are maxima and also using symmetry considerations when minima; its form will, naturally, be given in the limiting characterization section.

The definition of processes by a non-limiting technique is analogous to that of [Mann \(1953\)](#) for Brownian motion, using the know “duality”, for independence, between sums and maxima (characteristic functions versus distribution functions), now extended to stochastic processes with continuous time. We also obtain stochastic processes, connected with maxima, whose one-dimensional finite behaviour was the one described asymptotically by [Newell \(1962\)](#) for diffusion processes, like the Orstein-Uhlenbeck Brownian motion. The structure of the jump process is obtained as well as a natural representation.

Let us recall some results.

A general Gumbel random variable has the distribution function $\Lambda((x - \lambda)/\delta)$ where $\Lambda(z) = (\exp(-e^{-z}))$ is the distribution function of a reduced random variable ($\lambda = 0, \delta = 1$); we also have $\Lambda'(z) = e^{-z} \Lambda(z)$. The first moments are $\mu = \lambda + \gamma \delta$ ($\gamma = 0.57722 \dots$, Euler's constant), $\sigma^2 = \pi^2/6 \cdot \delta^2 = 1.64493 \dots \delta^2$, $\sigma = \pi/6 \cdot \delta = 1.28825 \dots \delta$, $\beta_1 = \mu_3/\sigma^2 = 1.13955$, and $\beta_2 = \mu_4/\sigma^4 = 5.4$; see [Gumbel \(1958\)](#).

An extreme random pair (Z_1, Z_2) with reduced Gumbel margins has the distribution function $\Lambda(x, y) = \text{Prob}\{Z_1 \leq x, Z_2 \leq y\} = \exp\{-(e^{-x} + e^{-y}) k(y - x)\} = \{\Lambda(x)\Lambda(y)\}^{k(y-x)}$ where the dependence function $k(\cdot)$ satisfies some conditions for $\Lambda(\cdot, \cdot)$ to be a distribution function with reduced

Gumbel margins; see [Tiago de Oliveira \(1962/63\)](#). In particular we have $\frac{\max(1, e^w)}{1+e^w} \leq k(w) \leq 1$, the lower bound corresponding to the diagonal case where we have $\text{Prob}\{Y = X\} = 1$ (for reduced margins) and $k(w) = 1$ corresponding to independence.

The correlation coefficient has the expression

$$\rho = -\frac{6}{\pi^2} \int_{-\infty}^{+\infty} \log k(w) \, dw \quad (\geq 0),$$

$\rho = 0$ being equivalent to independence and $\rho = 1$ to the diagonal case.

The probability $D(w) = \text{Prob}\{Z_2 - Z_1 \leq w\}$, (Z_1, Z_2) being reduced Gumbel random variables, has the expression $D(w) = \frac{k'(w)}{k(w)} + \frac{e^w}{1+e^w}$ a.e., as $k'(\cdot)$ exists a.e.; in the independence case ($k(w) = 1$) we have the logistic distribution and in the diagonal case we have $D(w) = H(w)$, where $H(w)$ is the Heaviside jump function ($H(w) = 0$ if $w < 0$ and $H(w) = 1$ if $w \geq 0$). Note that $\partial^2 \Lambda / \partial x \partial y$ does not exist in all cases — as $k''(\cdot)$ does not always exist; this is the case in all stochastic processes (sequences) connected with extremes, as happens in extremal sequences and processes and in EMS and EME sequences.

In particular, for $k_\theta(w) = 1 - \frac{\min(\theta, e^w)}{1+e^w}$ ($0 \leq \theta \leq 1$), we have $D_\theta(w) = 0$ if $w < \log \theta$ (< 0) (i.e., $\text{Prob}\{Z_2 \geq Z_1 + \log \theta\} = 1$) and $D_\theta(w) = \frac{1}{1+(1-\theta)e^w}$ if $w \geq \log \theta$ with a jump of θ at $w = \log \theta$, the jump being the probability of $Y = X + \log \theta$; the correlation coefficient is

$$\rho(\theta) = \frac{6}{\pi^2} R(\theta)$$

where $R(\theta) = \int_0^{\theta - \log t} \frac{1}{1-t} \, dt$, $R'(\theta) \geq 0$ with $R(0) = 0$, $R(1) = \frac{6}{\pi^2}$ so that for $\theta = 0$ and $\theta = 1$ we have, obviously, the independence and diagonal cases. The expression of $k_\theta(w)$ corresponds to the *biextremal model*.

The other coefficients are the difference-sign correlation $\tau(\theta) = \theta$, the grade correlation $\chi(\theta) = 3\theta/(2 + \theta)$, and the medial correlation $v(\theta) = 2^\theta - 1$, all increasing from 0 to 1 as θ increases from $\theta = 0$ to $\theta = 1$.

14.2 Extremal sequences

Let $E_1, E_2, \dots, E_n, \dots$ be a sequence of i.i.d. reduced Gumbel random variables and let $Z_1 = E_1, Z_2 = \max(Z_1, E_2), \dots, Z_n = \max(Z_{n-1}, E_n) = \max(E_1, \dots, E_n)$.

Evidently $\text{Prob}\{Z_1 \leq z_1, \dots, Z_n \leq z_n\} = \text{Prob}\{Z_i \leq z_i, i = 1, \dots, n\} = \text{Prob}\{E_1 \leq \min_{1 \leq i \leq n} z_i, E_2 \leq \min_{2 \leq i \leq n} z_i, \dots\} = \prod_{i=1}^n \Lambda(\min(z_i, \dots, z_n))$. This

corresponds to putting $t_1 = 1, \dots, t_n = n$ in the extremal processes, so we will only sketch some results. We can now consider the parameterized sequence $X_n = \lambda + \delta Z_n$ with real λ and $\delta(> 0)$. We have $\text{Prob}\{X_i \leq x_1, \dots, X_n \leq x_n\} = \text{Prob}\{Z_i \leq (x_i - \lambda)/\delta, \dots, Z_n \leq (x_n - \lambda)/\delta\} = \prod_{i=1}^n \Lambda((\min(x_i, \dots, x_n) - \lambda)/\delta)$.

Let us now consider a periodic subsequence $X'_j = X_{jp}$, $p(> 0)$ integer. We can show, in two ways, that $\{X'_j\}$ and $\{X_j\}$ have the same distribution apart from the parameters, which are (λ, δ) for the sequence $\{X_n\}$ and $(\lambda', \delta') = (\lambda + \delta \log p, \delta)$ for the sequence $\{X'_n\}$. A simple one is to consider the sequence $E'_1 = \max(E_1, \dots, E_p) - \log p, E'_2 = \max(E_{p+1}, \dots, E_{2p}) - \log p$, etc. It is immediate that the E'_j are independent random variables with the reduced Gumbel distribution. If we denote by $\{Z'_i = E'_1, Z'_2 = \max(E'_1, E'_2), \dots, Z'_n = \max(E'_1, \dots, E'_n)\}$, we see that

$$Z'_j = Z_{jp} - \log p \quad \text{and so} \quad X'_j = \lambda + \delta(Z'_j + \log p) = \lambda' + \delta Z'_j,$$

and then $\text{Prob}\{X'_j \leq x_1, \dots, X'_n \leq x_n\} = \prod_{i=1}^n \Lambda((\min(x_i, \dots, x_n) - \lambda')/\delta)$,

which is the same as that of $\{X_j\}$ with the new parameters.

Thus a change of timescale — or a periodic selection at instants $p, 2p, \dots$ — leads to extremal sequences with the same properties. In particular we have $M(X_j) = \lambda + \delta(\gamma + \log j), C(X_j, X_l) = \delta^2 R(\frac{\min(j,l)}{\max(j,l)}), M(X'_j) = M(X_{jp}) = \lambda + \delta(\gamma + \log(jp)) = \lambda' + \delta \log j$, and $C(X'_j, X'_l) = \delta^2 R(\frac{\min(j,l)}{\max(j,l)})$.

Periodic sub-sampling does not alter the structure of the $\{X_j\}$ apart from the location effect of the period.

Another proof is analogous to the one given for extremal processes where this effect is integrated in a general change of the time scale.

An ergodic (type) theorem, analogous to the one for extremal processes, can be written. It is $b/a \rightarrow \infty$,

$$\frac{1}{\log(b/a)} \sum_{k=1}^b \frac{X_k}{k} - \delta \log \sqrt{a b} \xrightarrow{ms} \lambda + \gamma \delta.$$

The proof is not important being analogous to the one to follow and statistically irrelevant.

14.3 Extremal processes: distribution and properties

A stochastic process $Z(t)$ ($t \geq 0$) is called an *extremal process* (of maxima) if:

- 1) $Z(t + \theta) = \max(Z(t), E(t, \theta))$, where $X(t)$ and $E(t, \theta)$ are independent random variables ($\theta \geq 0$);
- 2) $E(t, \theta)$ and $E(t', \theta')$ are independent random variables if the time intervals $]t, t + \theta]$ and $]t', t' + \theta']$ are disjoint;
- 3) The processes is time-homogeneous, that is, the distribution of $E(t, \theta)$ depends only on θ and is independent of t ;
- 4) $X(0) = -\infty$.

If $G(x)$ is the distribution function of $X(1)$ we have

$$\text{Prob} \{Z(t) \leq x\} = G^t(x) \text{ and } \text{Prob} \{E(t, \theta) \leq x\} = G^\theta(x).$$

Let $F(x, t)$ be the distribution function of $X(t)$ and denote by $F_1(x, \theta)$ the distribution function of $E(t, \theta)$, independent of t by assumption (3).

Thus we have, by (1) and (2),

$$\begin{aligned} F(x, t + \theta + \theta') &= F(x, t) F_1(x, \theta + \theta') = F(x, t + \theta) F_1(x, \theta') \\ &= F(x, t) F_1(x, \theta) F_1(x, \theta') \end{aligned}$$

so that

$$F_1(x, \theta + \theta') = F_1(x, \theta) F_1(x, \theta')$$

As the solution of this functional equation, for measurable functions, is the exponential function, we have

$$F_1(x, \theta) = F_1^\theta(x, 1)$$

As $F(x, t) = F(x, 0) F_1(x, t) = F(x, 0) F_1^t(x, 1)$ and by (4). $F(x, 0) = 1$ we obtain

$$F(x, t) = F_1^t(x, 1) = G^t(x)$$

as

$$F(x, 1) = F_1(x, 1) = G(x).$$

This result was obtained asymptotically by [Newell \(1962\)](#) under other conditions.

Let us finally introduce a new assumption, that one-dimensional margins have Gumbel distributions, by:

5) $G(x) = \Lambda(x)$ (standardization of the time unit).

We have then

The one-dimensional distribution of $Z(t)$ is $\Lambda^t(x) = \Lambda(x - \log t)$;

the two-dimensional distribution of $(Z(t_1), Z(t_2))$ ($t_1 < t_2$) is

$$\Lambda(\min(x_1, x_2) - \log t_1) \cdot \Lambda(x_2 - \log(t_2 - t_1));$$

and, in general, the n -dimensional distribution of $(Z(t_1), Z(t_2), \dots, Z(t_n))$ is

$$\prod_{i=1}^n \Lambda(\min(x_1, \dots, x_n) - \log(t_i - t_{i-1})) (t_0 = 0 < t_1 < \dots < t_n).$$

It is also immediate that, with $H(w)$ denoting the Heaviside distribution function of the almost sure random variable equal to zero, then

$X(t)$ is a Markoff process with transition probability function

$$P(x_1, t_1; x_2, t_2) = H(x_2 - x_1) \Lambda^{t_2 - t_1}(x_2).$$

From this transition probability function we get, denoting by $q(x_1, t_1)$ and $Q(t_1, x_1, x_2)$ the *process intensity* and the (relative) *transition probability function* (using formula (8.8.5) of [Fisz \(1962\)](#)),

$$P(x_1, t_1; x_2, t_2) = [1 - q(x_1, t_1)(t_2 - t_1)] H(x_2 - x_1)$$

$$+ q(x_1, t_1) Q(t_1, x_1, x_2)(t_2 - t_1) + 0(t_2 - t_1),$$

the following result:

An extremal process is a purely discontinuous (step) Markoff process where

$$P(x_1, t_1; x_2, t_2) = [1 - e^{-x_1}(t_2 - t_1)] H(x_2 - x_1) + o(t_2 - t_1),$$

so that

$$q(x_1, t_1) = e^{-x_1} \quad \text{and} \quad Q(t_1, x_1, x_2) = H(x_2 - x_1)(1 - e^{-x_1 - x_2}).$$

For this purely discontinuous Markoff process we can immediately write the classical Kolmogoroff-Chapman integral equations.

Let $T(a)$ denote the *upcrossing time* of the level a . As $\text{Prob}\{T(a) > t\} = \text{Prob}\{Z(t') < a, t' \leq t\} = \Lambda^t(a) = \Lambda(a - \log t)$, we can say that:

The distribution function of the upcrossing time is

$$\text{Prob}\{T(a) \leq t\} = 1 - \Lambda(a - \log t).$$

Some more results can also be given.

A first one, relating the chance experiments between points of time, is that for $s \leq u \leq t$ we have $E(s, t) = \max(E(s, u), E(u, t))$, as follows from the two expressions of $Z(t)$ in $Z(s)$, directly and by way of $Z(u)$.

As $Z(t) - \log t = E(0, t) - \log t$ is a reduced Gumbel random variable we see that $\frac{Z(t)}{\log t} \xrightarrow{ms} 1$.

Finally the least squares predictor of $Z(t)$, knowing $Z(t_1) = z_1, \dots, Z(t_n) = z_n$ ($0 < t_1 < t_2 < \dots < t_n < t$), as $Z(t)$ is markovian, it depends only on z_n is given by

$$p(t; z_n, t_n) = z_n + \int_{z_n}^{+\infty} (1 - \Lambda(z - \log(t - t_n))) dz \quad (\geq z_n)$$

whose mean-square error is

$$\int_{-\infty}^{+\infty} [\int_{z_n}^{+\infty} (1 - \Lambda(y - \log(t - t_n))) dy^2 - [p^2(t; z_n, t_n) - z_n^2] d\Lambda^{t-t_n}(z_n)].$$

From the practical point of this is not very rewarding and it is simpler to use the linear predictor $Z(t) = z_n + \log t/t_n$.

Notice that, as $Z(t)$ is not decreasing, we know that $\tilde{Z}(t) = \max_{0 \leq s \leq t} Z(s) = Z(t)$.

14.4 The limiting characterizations

We will sketch another characterization of extremal processes; for details and enlargements see Dwass (1964, 1966, 1973), Lamperti (1964), Resnick (1973, 1975) and Resnick and Rubinovitch (1973); for other approaches see references in Galambos (1978) and Leadbetter, Lindgren and Rootzén (1983).

Let $\{X_j\}, j = 1, 2, \dots$ be a sequence of i.i.d. random variables with distribution function $F(x)$ and such that there exist λ_n and $\delta_n (> 0)$ such that

$$\text{Prob} \left\{ \max_{1 \leq i \leq n} (X_i - \lambda_n) / \delta_n \leq x \right\} = F^n(\lambda_n + \delta_n x) \xrightarrow{w} L(x),$$

where the proper and non-degenerate limiting distribution is necessarily a Gumbel distribution function $\Lambda(x) = \exp(-e^{-x})$, or a Fréchet distribution function $\Phi_\alpha(x) = 0$ if $x \leq 0$, $\Phi_\alpha(x) = \exp(-e^{-\alpha x})$ if $x \geq 0$ with $\alpha > 0$, or even a Weibull distribution function $\Psi_\alpha(x) = \exp(-(-x)^\alpha)$ if $x \leq 0$, $\Psi_\alpha(x) = 1$ if $x \geq 0$ with $\alpha > 0$.

Let us define, for each $n = 1, 2, \dots$ and $t \geq 0$, stochastic processes

$$M_n(t) = (\max(X_1, \dots, X_{[nt]}) - \lambda_n) / \gamma_n \text{ for } t > 1/n$$

$$= (X_1 - \lambda_n) / \delta_n \text{ for } 0 \leq t \leq 1/n,$$

$[u]$ denoting as usual the largest integer $\leq u$.

$$\text{As } \text{Prob}\{M_n(t) \leq x\} = F^{[nt]}(\lambda_n + \delta_n x) = [F^n(\lambda_n + \delta_n x)]^{[nt]/n} \rightarrow L^t(x)$$

and more generally, for $0 < t_1 < t_2 < \dots < t_k$,

$$\text{Prob}\{M_n(t_1) \leq x_1, M_n(t_2) \leq x_2, \dots, M_n(t_k) \leq x_k\} \rightarrow$$

$$L^{t_1}(\min_{1 \leq i \leq k} x_i) L^{t_2 - t_1}(\min_{2 \leq i \leq k} x_i) L^{t_3 - t_2}(\min_{3 \leq i \leq k} x_i) \dots L^{t_k - t_{k-1}}(x_k),$$

we define an extremal reduced process $M(t)$ as a limit of $M_n(t)$ which extends to Fréchet and Weibull margins the extremal process defined with Gumbel margins.

It is evident that all “margin-free” properties such as the jump process described below. (to be Poisson-distributed) and all non-parametric correlation coefficients have the same behaviour as in extremal process with Gumbel

margins, which will continue to be considered in what follows. Note, in particular, that second order properties are not always valid for extremal processes with Fréchet or Weibull margins.

14.5 Second order properties of extremal processes

The second order properties of extremal processes (with Gumbel margins, to be dealt with in the rest of the chapter) depend on the form of the covariance (or correlation) function, as the margins have mean value and variance.

From the expression of the two-dimensional distribution function of $(Z(t_1), Z(t_2))$ passing to the reduced margins $\xi_1 = x_1 - \log t_1$ and $\xi_2 = x_2 - \log t_2$, we see that the mean value function of $Z(t)$ is $\mu(t) = \gamma + \log t$ and the dependence function ($t_1 \leq t_2$) is the biextremal one

$$k(w; t_1/t_2) = \frac{(1-t_1/t_2) + \max(e^w, t_1/t_2)}{1+e^w} = 1 - \frac{\min(t_1/t_2, e^w)}{1+e^w},$$

with parameter $\theta = t_1/t_2$ ($0 \leq \theta \leq 1$). Consequently:

An extremal (Gumbel) location-dispersion free process has the mean value $\gamma + \log t$, variance $\pi^2/6$ and the correlation function

$$\rho(t_1, t_2) = \rho(t_1/t_2) = 1 + \frac{6}{\pi^2} \int_{t_1/t_2}^1 \frac{\log \alpha}{1-\alpha} d\alpha = -\frac{6}{\pi^2} \int_0^{t_1/t_2} \frac{\log \alpha}{1-\alpha} d\alpha =$$

$$\frac{6}{\pi^2} R(t_1/t_2)$$

the processes is integrable, continuous and indefinitely differentiable in mean square except at origin (0,0) where ρ is not continuous.

The mean-square continuity is, evidently, consistent with the almost sure discontinuity, the discontinuities being moving (random) discontinuities; the graph of the process is a non-decreasing random step function.

14.6 The jumps of an extremal process

As we showed, the transition probability function is

$$P(x_1, t_1; x_2, t_2) = H(x_2 - x_1) \Lambda^{t_2-t_1}(x_2)(t_1 < t_2).$$

Consequently the probability of zero jumps in the time interval $]t_1, t_2]$ is given by

$$p_0(t_1, t_2) = \int_{-\infty}^{+\infty} \Lambda^{t_2-t_1}(x) d\Lambda^{t_1}(x) = t_1/t_2,$$

because the probability of zero jumps in $]t_1, t_2]$ if $Z(t_1) = x_1$ (which implies $Z(t_2) = x_1$) is $\Lambda^{t_2-t_1}(x_1)$; the probability that we have at least one jump is then

$$p_{>0}(t_1, t_2) = 1 - t_1/t_2.$$

Notice that for every $t > 0$ we have $\text{Prob}\{Z(t) > -\infty\} = p_{>0}(0, t) = 1$ so that:

For $t > 0$, $Z(t)$ is finite with probability one.

If $PN(t_1, t_2)$ denotes the probability of N jumps in the interval $]t_1, t_2]$, the probability of zero jumps in the interval $]t_2, t_2 + \theta]$ is given by

$$p_0(t_2, t_2 + \theta) = \frac{t_2}{t_2 + \theta} = 1 - \frac{\theta}{t_2} + o(\theta),$$

and as

$$p_{>0}(t_2, t_2 + \theta) = \frac{\theta}{t_2} + o(\theta),$$

as is natural and will be proved later, then

$$p_1(t_2, t_2 + \theta) = p_{>0}(t_2, t_2 + \theta) + o(\theta).$$

The basic equation is

$$\begin{aligned} PN + 1(t_1, t_2 + \theta) &= PN + 1(t_1, t_2) \cdot p_0(t_2, t_2 + \theta) \\ &+ PN(t_1, t_2) \cdot p_1(t_2, t_2 + \theta) + o(\theta). \end{aligned}$$

which, as derivatives exist, can be written as

$$PN + 1(t_1, t_2) + \frac{\partial PN + 1(t_1, t_2)}{\partial t_2} \theta = PN + 1(t_1, t_2) \left(1 - \frac{\theta}{t_2}\right) + PN(t_1, t_2) \frac{\theta}{t_2} + o(\theta).$$

and gives, for $\theta \rightarrow 0$,

$$\frac{\partial PN + 1(t_1, t_2)}{\partial t_2} = \frac{1}{t_2} (PN(t_1, t_2) - PN + 1(t_1, t_2)).$$

If we denote by $\Delta(z; t_1, t_2) = \sum_0^\infty PN(t_1, t_2) z^N$ the generating function of the probabilities, then $\Delta(1; t_1, t_2) = 1$ and $\Delta(z; t_1, t_1) = 1$.

The generating function Δ , from the difference-differential equation for the probabilities, satisfies the partial differential equation

$$\frac{1}{z} \left(\frac{\partial \Delta}{\partial t_2} - \frac{\partial p_0(t_1, t_2)}{\partial t_2} \right) = \frac{1}{t_2} \left(\Delta \frac{\Delta - p_0(t_1, t_2)}{z} \right).$$

From $p_0(t_1, t_2) = t_1/t_2$ we obtain

$$\frac{1}{\Delta} \frac{\partial \Delta}{\partial t_2} = \frac{z-1}{t_2},$$

whose solution is evidently

$$\Delta(z; t_1, t_2) = \left(\frac{t_2}{t_1} \right)^{z-1} = e^{(z-1) \log(t_2/t_1)}$$

owing to the known values of $\Delta(1; t_1, t_2) = 1$ and $\Delta(z; t_1, t_1) = 1$. But Δ is the generating function of the Poisson distribution with parameter $v = \log(t_2/t_1)$, and so

The jumps of $Z(t)$ in the interval, $[t_1, t_2]$ have a Poisson distribution with parameter $v = \log(t_2/t_1)(> 0)$; the mean number and variance of jumps increase logarithmically with time.

Thus in the finite interval $]0, t]$ we must expect infinity of jumps; the natural science interpretation is easy: in the process beginning at $-\infty (Z(0) = -\infty)$ we should expect an infinite number of jumps to obtain any finite value.

Let us show, finally, that

$p_1(t, t + \theta) = \frac{\theta}{t} + 0(\theta)$; as $p_{>0}(t, t + \theta) = \frac{\theta}{t} + 0(\theta)$, it is sufficient to show that $p_{>1}(t, t + \theta) = 0(\theta)$. As we have two or more jumps on $]t, t + \theta]$, let $\tilde{\theta}$ denote the mid-point of the instants of the first and last jump and $\pi(\tilde{\theta})$ its distribution function. We can clearly write

$$\begin{aligned} p_{>1}(t, t + \theta) &= \int_0^\theta p_{>0}(t, t + \tilde{\theta}) p_{>0}(t + \tilde{\theta}, t + \theta) d\pi(\tilde{\theta}) \\ &= \int_0^\theta \frac{\tilde{\theta}}{t + \tilde{\theta}} \cdot \frac{\theta - \tilde{\theta}}{t + \theta} d\pi(\tilde{\theta}) \leq \frac{\theta^2}{t^2} \int_0^\theta \frac{\tilde{\theta}}{\theta} \cdot \frac{\theta - \tilde{\theta}}{\theta} d\pi(\tilde{\theta}) \leq \frac{\theta^2}{t^2} \end{aligned}$$

with $p_{>1}(t, t + \theta) = 0(\theta)$ as we wished to show.

The following representation (i.e., a stochastic process with the same finite-dimensional distributions) can be given:

For $t \geq t_0$ the extremal process $Z(t)$ is equivalent in law to

$$Z(t) = Z(t_0) \text{ for } t \in [t_0, t_1[,$$

$$Z(t) = Z(t_1) \text{ for } t \in [t_1, t_2[$$

...

where $t_1 < t_2 < \dots$ are the jump times after t_0 , having the densities t_0/t_1^2 for $t_1 \geq t_0$, t_1/t_2^2 for $t_2 \geq t_1$, ... and $Z(t_1) = x_1, Z(t_2) = x_2, \dots$ are the values at the jump times $t_1 < t_2 < \dots$ with the distribution functions

$$\frac{\Lambda^{t_1-t_0}(x_1) - \Lambda^{t_1-t_0}(x_0)}{1 - \Lambda^{t_1-t_0}(x_0)} \text{ for } x_1 > x_0,$$

$$\frac{\Lambda^{t_2-t_1}(x_2) - \Lambda^{t_2-t_1}(x_1)}{1 - \Lambda^{t_2-t_1}(x_1)} \text{ for } x_2 > x_1,$$

...

The proof is immediate because the probability of no jumps in the time interval $[t_0, t_1[$ is t_0/t_1 and the probability that, as there is a jump, its value is smaller than x_1 is given by the transition probability function truncated at x_0 .

The study of the times of the jumps, as will be seen, is not helpful for statistical decision making.

14.7 The logarithmization of time and the ergodic theorem for extremal processes

As $\log t$ appears frequently in the expressions connected with extremal processes, we can consider the use of a new time variable $\tilde{t} = \log t$ (logarithmic time) defining a new process $\tilde{Z}(\tilde{t}) = Z(e^{\tilde{t}})$, whose characterization could be transposed from that of $Z(t)$. We will give a different one, equivalent but more natural for the case. Let us just remark that $\text{Prob} \{\tilde{Z}(\tilde{t}) \leq x\} = \Lambda(x - \tilde{t})$.

Consider then the (transposed) extremal process:

- 1) $\tilde{Z}(\tilde{t})$ is defined over all the real (time) line;

2) $\tilde{Z}(\tilde{t}_2) = \max(\tilde{Z}(\tilde{t}_1), \tilde{E}(\tilde{t}_1, \tilde{t}_2))$ where $\tilde{Z}(\tilde{t}_1)$ and $\tilde{E}(\tilde{t}_1, \tilde{t}_2)$ are independent random variables;

3) $\tilde{Z}(\tilde{t})$ has the Gumbel distribution $\Lambda(x - \tilde{t})$ for convenient standardization of the time (or change of origin) and change of scale (i.e., a linear change of the clock).

Note immediately, that from (3) we have $\tilde{Z}(-\infty) = -\infty$ with probability one.

Let us denote by $\Psi(x_1, \tilde{t}_1; x_2, \tilde{t}_2) = \text{Prob}\{\tilde{Z}(\tilde{t}_1) \leq x_1, \tilde{Z}(\tilde{t}_2) \leq x_2\}$; we have, $\Psi(x_1, \tilde{t}_1; x_2, \tilde{t}_2) = \text{Prob}\{\tilde{Z}(\tilde{t}_1) \leq x_1, \tilde{Z}(\tilde{t}_2) \leq x_2, \tilde{E}(\tilde{t}_1, \tilde{t}_2) \leq x_2\} = \Lambda(\min(x_1, x_2) - \tilde{t}_1) \cdot G(x_2; \tilde{t}_1, \tilde{t}_2)$ using hypothesis (3) for the distribution of $\tilde{Z}(\tilde{t}_1)$.

With hypothesis (3) we have also, putting $x_1 = +\infty$,

$$\Lambda(x_2, \tilde{t}_2) = \Lambda(x_2, \tilde{t}_1) \cdot G(x_2; \tilde{t}_1, \tilde{t}_2)$$

and so

$\Psi(x_1, \tilde{t}_1; x_2, \tilde{t}_2) = \Lambda(\min(x_1, x_2) - \tilde{E}_1) \Lambda(x_2 - \log(e^{\tilde{t}_2} - e^{\tilde{t}_1}))$, as could be expected from the relation $t = e^{\tilde{t}}$.

Thus the mean value function is $\gamma + \tilde{t}$, the variance $\frac{\pi^2}{6}$ and the correlation function

$$\tilde{\rho}(\tilde{t}_2 - \tilde{t}_1) = \rho(e^{\tilde{t}_2}/e^{\tilde{t}_1}) = \frac{6}{\pi^2} R(e^{-|\tilde{t}_2 - \tilde{t}_1|}),$$

so $\tilde{Y}(\tilde{t}) = \tilde{Z}(\tilde{t}) - \gamma - \tilde{t}$ is a weakly stationary process.

$\tilde{Y}(\tilde{t})$ is thus an ergodic process (of second order) and we know that $\frac{1}{b-a} \int_a^b \tilde{Y}(\tilde{t}) d\tilde{t} \xrightarrow{ms} 0$ as $b - a \rightarrow +\infty$.

Passing to the original clock $t = e^{\tilde{t}}$, $b = e^{\tilde{b}}$, $a = e^{\tilde{a}}$, we get

$$\frac{1}{\log(b/a)} \int_a^b \frac{Z(t) - \gamma - \log t}{t} dt \xrightarrow{ms} 0 \text{ as } b/a \rightarrow +\infty$$

or, in a simpler form,

$$\frac{1}{\log(b/a)} \int_a^b \frac{Z(t)}{t} dt - \log \sqrt{a/b} \xrightarrow{ms} \gamma$$

when $b/a \rightarrow +\infty$. We have for the variance of $\frac{1}{\log(b/a)} \int_a^b \frac{Z(t)}{t} dt$ the expression

$$R(a/b) + \frac{2}{(\log(b/a))^2} \int_{a/b}^1 \left(\log \frac{b}{a} + \frac{\log u}{2} \right) \frac{(\log u)^2}{1-u} du.$$

Notice that if the margins are Fréchet or Weibull, the correlation coefficient between $Z(t_1)$ and $Z(t_2)$ ($0 < t_1 < t_2$) also depends only on t_1/t_2 (and α), when the second moments exist.

14.8 The parameterization of extremal processes

Let $Z(t), t \geq 0$ be a (reduced) extremal process and consider the two (general) extremal processes

$$X_1(t) = \lambda + \delta Z(t) \quad (\lambda \text{ real}, \delta > 0) \text{ and}$$

$$X_2(t) = \delta Z(\beta t) \quad (\delta, \beta > 0).$$

Their mean values $\mu_1(t) = \lambda + \delta(\gamma + \log t)$ and $\mu_2(t) = \delta(\gamma + \log(\beta t))$ are equal if $\lambda = \delta \log \beta$; also the covariance function in both processes is $R(\frac{\min(s,t)}{\max(s,t)})\delta^2$. This suggests that $X_1(t)$ and $X_2(t)$ can have identical distribution. In fact we have (with $t_0 = 0$)

$$\text{Prob}\{X_1(t_i) \leq x_i\} = \text{Prob}\{Z(t_i) \leq (x_i - \lambda)/\delta\}$$

$$= \prod_{i=1}^n \Lambda\left(\min_{j=i}^n \frac{x_j - \lambda}{\delta}\right) - \log(t_i - t_{i-1})$$

$$= \prod_{i=1}^n \Lambda\left\{\left(\min_{j=i}^n x_j - \lambda\right)/\delta - \log(t_1 - t_{i-1})\right\}$$

and also

$$\text{Prob}\{X_2(t_i) \leq x_i\} = \text{Prob}\{Z(\beta t_i) \leq x_i/\delta\}$$

$$= \prod_{i=1}^n \Lambda\left(\min_{j=i}^n \frac{x_j}{\delta} - \log(\beta(t_i - t_{i-1}))\right)$$

$$= \prod_{i=1}^n \Lambda((\min_{j=i}^n x_j - \delta \log \beta) / \delta - \log(t_i - t_{i-1}))$$

which are equal if $\lambda = \delta \log \beta$ as said before.

Evidently any splitting (λ_1, β_1) with $\lambda = \delta \log \beta = \lambda_1 + \delta \log \beta_1$ is admissible.

Owing to linearity we will prefer the parameterization in (λ, δ) ; any change of time unit (scale) is given by $\lambda = \delta \log \beta$.

14.9 Statistics for extremal sequences

Let $\{Z_n, n \geq 1\}$ be a extremal sequence, as described before, with mean values $\mu_i = M(Z_i) = \gamma + \log i$ and covariance function $\sigma_{ij} = (Z_i, Z_j) = R(\frac{\min(i,j)}{\max(i,j)})$ with $R(u) = -\int_0^u \frac{\log t}{1-t} dt$ ($R(0) = 0, R'(u) > 0, R(1) = \pi^2/6$), and consider the sequence $\{X_n, n \geq 1\}$, with $X_n = \lambda + \delta Z_n$, λ real and $\delta > 0$; as said before we can have a subsequence $\{X_{np}\}$ of $\{X_n\}$ with $p(> 0)$ an integer, obtained from the original by periodic sampling, with the convenient change of parameters (the location one, in this case). We know then that $M(X_i) = \lambda + \delta \mu_i$ and $C(X_i, X_j) = \delta^2 \sigma_{ij} = \delta^2 R(\frac{\min(i,j)}{\max(i,j)})$.

As $X_i = \lambda + \delta Z_i = \lambda + \delta \mu_i + \varepsilon_i$ with $\varepsilon_i = \delta(Z_i - \mu_i)$ where $M(\varepsilon_i) = 0$ and $C(\varepsilon_i, \varepsilon_j) = \delta^2 \sigma_{ij} = \delta^2 R(\frac{\min(i,j)}{\max(i,j)})$, we know that the least-squares (unbiased) estimators of (λ, δ) are given by

$$\begin{bmatrix} \lambda^* \\ \delta^* \end{bmatrix} = (A^T \Sigma^{-1} A)^{-1} A^T \Sigma^{-1} [X]$$

where $A^T = \begin{bmatrix} 1 & \dots & 1 \\ \mu_1 & \dots & \mu_n \end{bmatrix}$, $\Sigma = \delta^2 [\sigma_{ij}]$ (variance-covariance matrix), $[X]^T = [X_1, \dots, X_n]$ (the transposed sample vector) and $\det(A^T \Sigma^{-1} A)$ assumed to be $\neq 0$; see [Silvey \(1975\)](#). The least squares estimator of the quantile $\varphi = \lambda + \chi \delta$ is, as known, $\varphi^* = \lambda^* + \chi \delta^*$, which is unbiased ($M(\varphi^*) = \varphi$) and has the variance $V(\varphi^*) = \delta^2 [1 \chi] (A^T \Sigma^{-1} A)^{-1} \begin{bmatrix} 1 \\ \chi \end{bmatrix} = V(\lambda^*) + 2\chi C(\lambda^*, \delta^*) + \chi^2 V(\delta^*)$ with $(A^T \Sigma^{-1} A)^{-1} = \begin{bmatrix} V(\lambda^*) & C(\lambda^*, \delta^*) \\ C(\lambda^*, \delta^*) & V(\delta^*) \end{bmatrix}$.

From the form of $\Sigma = [R(\min(i,j)/\max(i,j))]$, the least squares estimators (λ^*, δ^*) do not seem very easy to manage.

Another pair of (unbiased) estimators is given by minimizing

$$\Sigma_1^n \in \frac{2}{i} = [X - \lambda - \delta \mu]^T [X - \lambda - \delta \mu] \quad \text{as}$$

$$\begin{bmatrix} \lambda^{**} \\ \delta^{**} \end{bmatrix} = (A^T A)^{-1} A^T [X] = \begin{bmatrix} n & \sum_i \mu_i \\ \sum_i \mu_i & \sum_i \mu_i^2 \end{bmatrix}^{-1} \begin{bmatrix} 1 & \dots & 1 \\ \mu_1 & \dots & \mu_n \end{bmatrix} \begin{bmatrix} X_1 \\ \dots \\ X_n \end{bmatrix}$$

$$= \begin{bmatrix} n & \sum_i \mu_i \\ \sum_i \mu_i & \sum_i \mu_i^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_i X_i \\ \sum_i \mu_i X_i \end{bmatrix}$$

$$= \frac{1}{n \sum_i \mu_i^2 - (\sum_i \mu_i)^2} \begin{bmatrix} \sum_i \mu_i^2 \cdot \sum_i X_i - \sum_i \mu_i \cdot \sum_i \mu_i X_i \\ - \sum_i \mu_i \cdot \sum_i X_i + n \sum_i \mu_i X_i \end{bmatrix}.$$

The estimator of φ is, then, $\varphi^{**} = \lambda^{**} + \chi \delta^{**}$ and

$$V(\varphi^{**}) = [1 \chi] (A^T A)^{-1} (A^T \Sigma A) (A^T A)^{-1} \begin{bmatrix} 1 \\ \chi \end{bmatrix}.$$

The efficiency of φ^{**} with respect to φ^* is given by

$$\frac{V(\lambda^*) + 2 \chi C(\lambda^*, \delta^*) + \chi^2 V(\delta^*)}{V(\lambda^{**}) + 2 \chi C(\lambda^{**}, \delta^{**}) + \chi^2 V(\delta^{**})}$$

where

$$\begin{bmatrix} V(\lambda^{**}) & C(\lambda^{**}, \delta^{**}) \\ C(\lambda^{**}, \delta^{**}) & V(\delta^{**}) \end{bmatrix} = (A^T A)^{-1} (A^T \Sigma A) (A^T A)^{-1}.$$

The general efficiency of the system $(\lambda^{**}, \delta^{**})$ with respect to (λ^*, δ^*) can be defined in two ways:

a) as the value of

$$\frac{\det(A^T \Sigma^{-1} A)^{-1}}{\det\{(A^T A)^{-1} (A^T \Sigma A) (A^T A)^{-1}\}} = \frac{\det^2(A^T A)}{\det(A^T \Sigma A) \det(A^T \Sigma^{-1} A)};$$

see [Cramér \(1946\)](#);

b) as the smallest root of

$$\det \left(\begin{bmatrix} V(\lambda^*) & C(\lambda^*, \delta^*) \\ C(\lambda^*, \delta^*) & V(\delta^*) \end{bmatrix} - \xi \begin{bmatrix} V(\lambda^{**}) & C(\lambda^{**}, \delta^{**}) \\ C(\lambda^{**}, \delta^{**}) & V(\delta^{**}) \end{bmatrix} \right) = 0;$$

see [Tiago de Oliveira \(1972\) and \(1982\)](#).

Both definitions impose a lot of numerical calculation, beginning with the inversion of $\Sigma = [R(\min(i, j)/\max(i, j))]$, and do not seem easy to manipulate. If $n = 2$ we get for both estimators.

$$\begin{bmatrix} \lambda^* \\ \delta^* \end{bmatrix} = \begin{bmatrix} \lambda^{**} \\ \delta^{**} \end{bmatrix} = \frac{1}{\log 2} \begin{bmatrix} (\gamma + \log 2) X_1 - \gamma X_2 \\ X_2 - X_1 \end{bmatrix}$$

and both efficiencies are equal to 1.

In general we use the estimators:

$$\lambda^* = \frac{(\gamma \sum_i \log i + \sum_i (\log i)^2) \cdot \sum_i X_i - (n \gamma + \sum_i \log i) \cdot \sum_i (\log i) X_i}{n \sum_i (\log i)^2 - (\sum_i \log i)^2}$$

$$\delta^* = \frac{n \sum_i (\log i) X_i - \sum_i \log i \cdot \sum_i X_i}{n \sum_i (\log i)^2 - \sum_i (\log i)^2}.$$

whose variance-covariance matrix was given above.

14.10 Statistics for extremal processes

Consider now the extremal process $Z(t)$ ($t \geq 0$) ([Tiago de Oliveira, 1968](#)) whose mean value function is $\mu(t) = \gamma + \log t$, covariance function $\sigma(s, t) = C(Z(s), Z(t)) = R(\min(s, t)/\max(s, t))$ where, as before, $R(u) = -\int_0^u \frac{\log t}{1-t} dt$, $R(0) = 0$, $R'(0) > 0$, $R(1) = \pi^2/6$; the correlation function is $\rho(s, t) = \frac{6}{\pi^2} R(s, t)$.

Let us consider the extremal process $X(t) = \lambda + \delta Z(t)$, $t \geq 0$. It is well known (see above) that there is an essential *unidentifiability* problem between the use of $X(t)$, $X_1(t) = \lambda_1 + \delta Z(\beta_1 t)$ and $X_2(t) = \delta Z(\beta_2 t)$ with the equalities $\lambda = \lambda_1 + \delta \log \beta_1 = \delta \log \beta_2$. We have chosen the present formulation because then the parameters are linear. We will suppose that $X(t)$ is known (observed) for the interval $[a, b]$ or $]a, b[$, which is irrelevant, with $0 < a < b$.

We could approach the problem by an analogous procedure to the least squares, substituting the sums by integrals, i.e., as $X(t) = \lambda + \delta Z(t) = \lambda + \delta(\gamma + \log t) + \varepsilon(t)$ ($\varepsilon(t) = \delta(Z(t) - \gamma - \log t)$, with mean value zero), by minimizing

$$\int_a^b \int (X(t) - \lambda - \delta(\gamma + \log s)) \bar{\sigma}(s, t) (X(t) - \lambda - \delta(\gamma +$$

$\log t)) \, d s \, d t$ where $\bar{\sigma}(s, t)$ is the “inverse” of the covariance function $\sigma(s, t)$ analogous to the relation between Σ^{-1} and Σ . The solution is

$$\begin{bmatrix} \lambda^* \\ \delta^* \end{bmatrix} = \frac{1}{\bar{\sigma}_0 \bar{\sigma}_2 - \bar{\sigma}_1^2} \begin{bmatrix} \int_a^b \int (\bar{\sigma}_2 - \bar{\sigma}_1 \mu(t)) \bar{\sigma}(s, t) X(s) \, d s \, d t \\ \int_a^b \int (\bar{\sigma}_0 \mu(t) - \bar{\sigma}_1) \bar{\sigma}(s, t) X(s) \, d s \, d t \end{bmatrix}$$

where $\bar{\sigma}_0 = \int_a^b \int \bar{\sigma}(s, t) \, d s \, d t$, $\bar{\sigma}_1 = \int_a^b \int \bar{\sigma}(s, t) \mu(s) \, d s \, d t$ and $\bar{\sigma}_2 = \int_a^b \int \bar{\sigma}(s, t) \mu(s) \mu(t) \, d s \, d t$; (λ^*, δ^*) are unbiased estimators of (λ, δ) . The same result can be obtained by seeking the function $f(t)$ for which $\int_a^b f(t) X(t) \, d t$ is the least-squares (unbiased) estimator of $\varphi = \lambda + \chi \delta$, which can be shown, by a variational method, to be $\varphi^* = \lambda^* + \chi \delta^*$. But the result is unhelpful because the determination of $\bar{\sigma}(s, t)$ is practically impossible.

As before, for extremal sequences, it is natural to seek the estimators $(\lambda^{**}, \delta^{**})$ such that the (empirical) mean-square distance

$$\int_a^b (X(t) - \lambda - \delta \mu(t))^2 \, d t = \int_a^b (X(t) - \lambda - \delta(\gamma + \log t))^2 \, d t$$

is a minimum. We get the also unbiased estimators

$$\begin{bmatrix} \lambda^{**} \\ \delta^{**} \end{bmatrix} = \frac{1}{\int_a^b d t \cdot \int_a^b \mu^2(t) \, d t - (\int_a^b \mu(t) \, d t)^2} \begin{bmatrix} \int_a^b \mu^2(t) \, d t \cdot \int_a^b X(t) \, d t - \int_a^b \mu(t) \, d t \cdot \int_a^b \mu(t) X(t) \, d t \\ \int_a^b d t \cdot \int_a^b \mu(t) X(t) \, d t - \int_a^b \mu(t) \, d t \cdot \int_a^b X(t) \, d t \end{bmatrix}$$

Evidently we could choose two functions $l(t)$ and $d(t)$ and take $(\int_a^b l(t) X(t) \, d t, \int_a^b d(t) X(t) \, d t)$ as estimators of (λ, δ) with the condition of quasi-linearity or of unbiasedness; $(\lambda^{**}, \delta^{**})$, is one of these choices, optimization being practically impossible.

The complex structure of these estimators suggests the use of the ergodic theorem, when possible. As shown before,

$$\frac{1}{\log(b/a)} \int_a^b \frac{Z(t)}{t} \, d t - \log \sqrt{a b} \xrightarrow{m.s.} \gamma \text{ as } b/a \rightarrow +\infty$$

or, under a more general form,

$$\frac{1}{\log(b/a)} \int_a^b \frac{X(t)}{t} dt - \delta \log \sqrt{ab} \xrightarrow{ms} \lambda + \gamma \delta.$$

If we take $\log \sqrt{ab} = \chi - \gamma$ with $b/a \rightarrow +\infty$ or $a = e^{2(\chi-\gamma)}/b, b \rightarrow +\infty$ we see that

$$\frac{1}{2(\log b - (\chi - \gamma))} \int_{e^{2(\chi-\gamma)/b}}^b \frac{X(t)}{t} dt \xrightarrow{ms} \lambda + \chi \delta \text{ when } b \rightarrow +\infty.$$

As, then, $b/a = b^2/e^{2(\chi-\gamma)}$ increases with b , if we know $X(t)$ in an interval $[A, B]$ we can estimate $\varphi = \lambda + \chi \delta$ taking $b = B$ and $a = e^{2(\chi-\gamma)}/B$ if $a > A$. Thus the estimation is obtained by taking $\varphi^* = \frac{1}{2(\log B - (\chi - \gamma))} \int_{e^{2(\chi-\gamma)/B}}^b \frac{X(t)}{t} dt$ when $AB < e^{2(\chi-\gamma)}$ or $\chi \geq \gamma + \log \sqrt{AB}$, thus allowing, the estimators of large quantiles, which is important in the applications.

The variance of this estimator of $\lambda + \chi \delta$ is given by

$$\begin{aligned} \frac{V(\varphi^*)}{\delta^2} &= \frac{1}{4(\log B - (\chi - \gamma))^2} \int_{e^{2(\chi-\gamma)/B}}^B \int_{e^{2(\chi-\gamma)/B}}^B \frac{\bar{\sigma}(s,t)}{s t} ds dt \\ &= \frac{1}{2(\log B - \chi + \gamma)^2} \int_{e^{2(\chi-\gamma)/B}}^B \int_{e^{2(\chi-\gamma)/B}}^B R(s/t) \frac{ds dt}{st} \\ &= \frac{1}{2(\log B - \chi + \gamma)^2} \int_{e^{2(\chi-\gamma)/B^2}}^1 \frac{R(u)}{u} du \int_{e^{2(\chi-\gamma)/B}}^B \frac{dt}{t} \\ &= \frac{1}{2(\log B - \chi + \gamma)^2} \int_{e^{2(\chi-\gamma)/B^2}}^1 \frac{R(u)}{u} du \log \frac{B^2 u}{e^{2(\chi-\gamma)}} \\ &= \frac{1}{2(\log B - \chi + \gamma)^2} \int_{e^{2(\chi-\gamma)/B^2}}^1 \frac{R(u)}{u} (2(\log B - \chi + \gamma) + \log u) du \\ &= \frac{1}{2(\log B - \chi + \gamma)^2} \left\{ R(u) \left[2(\log B - \chi + \gamma) \log u + \frac{\log^2 u}{2} \right] \Big|_{e^{2(\chi-\gamma)/B^2}}^1 - \right. \\ &\quad \left. - \int_{e^{2(\chi-\gamma)/B^2}}^1 (2(\log B - \chi + \gamma) + \log u + \frac{\log^2 u}{2}) R'(u) du \right\} \\ &= \frac{1}{2(\log B - \chi + \gamma)^2} \left\{ (2(\log B - \chi + \gamma))^2 R\left(\frac{e^{2(\chi-\gamma)}}{B^2}\right) \right. \\ &\quad \left. - \int_{e^{2(\chi-\gamma)/B^2}}^1 (2(\log B - \chi + \gamma) + \log u + \frac{\log^2 u}{2}) R'(u) du \right\} = R(e^{2(\chi-\gamma)}/B^2) \end{aligned}$$

$$\begin{aligned}
 & - \frac{1}{2(\log B - \chi + \gamma)^2} \int_{e^{2(\chi-\gamma)/B^2}}^1 (2(\log B - \chi + \gamma) \log u + \frac{\log^2 u}{2}) \frac{\log u}{1-u} du \\
 & = R(e^{2(\chi-\gamma)/B^2}) - \frac{1}{2(\log B - \chi + \gamma)^2} \int_{e^{2(\chi-\gamma)/B^2}}^1 (2(\log B - \chi + \gamma) \log u + \frac{\log^2 u}{2}) \frac{\log u}{1-u} du \\
 & = R(e^{2(\chi-\gamma)/B^2}) - \int_{e^{2(\chi-\gamma)/B^2}}^1 \frac{\log^2 u}{1-u} du - \frac{1}{4(\log B - \chi + \gamma)^2} \int_{e^{2(\chi-\gamma)/B^2}}^1 \frac{\log^3 u}{1-u} du
 \end{aligned}$$

Analogously we could chose a point $C(0 < A < C < B)$ and study the stochastic integrals

$$\frac{1}{\log C/A} \int_A^{\chi(t)} \frac{\chi(t)}{t} dt \quad \text{and} \quad \frac{1}{\log B/C} \int_C^B \frac{\chi(t)}{t} dt$$

To have $\min(\log \frac{C}{A}, \log \frac{B}{C})$ the largest possible we must take $C = \sqrt{AB}$ with $\frac{C}{A} = \sqrt{B/A} = \frac{B}{C}$. The convex combination of those two integrals with coefficients $\frac{\log \sqrt{CB} - (\chi - \gamma)}{\log \sqrt{B/A}}$ and $\frac{(\chi - \gamma) - \log \sqrt{AC}}{\log \sqrt{B/A}}$ estimates, unbiasedly, and mean square, the quantile $\lambda + \chi \delta$ and consequently, for $\chi = \gamma + \log t, \lambda^* + (\gamma + \log t)\delta^*$ is the natural linear predictor of $\chi(t)$.

An important question in the integrals above is the assumed existence of jumps in the interval of observation $[A, B]$. Suppose that $\chi(t)$ has no jumps in $[A, B]$ a fact that has probability A/B . Noting by χ the value of $\chi(t)$ in the interval the estimator

$$(\int_A^B \ell(t) \chi(t) dt, \int_A^B d(t) \chi(t) dt) = (\chi \int_A^B \ell(t) dt, \chi \int_A^B d(t) dt) = (\chi, 0) (!) \quad \text{as by the quasilinearity we must have } \int_A^B \ell(t) dt = 1 \text{ and } \int_A^B d(t) dt = 0.$$

The probability of being capable to estimate is, then, $1 - A/B$.

In that case is easy to show that $\frac{\chi(B) - \chi(A)}{\log B/A} \xrightarrow{ms} \delta$ if $B/A \rightarrow +\infty$ but an analogous estimator of λ based in the extreme values $\chi(A)$ and $\chi(B)$ does not converge in m.s. to λ .

The study of the jump process does not see to illuminate the structure of the location parameter (λ) or of the change of time unit (β). In fact, the extremal process $Z(t)$ - and the processes $\lambda + \delta Z(t)$ or $\delta Z(\beta t)$ - have a random number of jumps $N(t_1, t_2)$ in the interval $[t_1, t_2](0 < t_1)$ with a Poisson distributions with parameter $(= \text{mean value} = \text{variance}) \log \frac{t_2}{t_1}$; in $[0, t]$ there is an infinity of jumps with probability one. If t_1 is a time instant (be it a jump instant or not) the

probability that the next jump instant is $< t_2$ is given by $1 - t_1/t_2 (t_1 < t_2)$. In any case a change of time unit does not alter the structure.

The mean value of inter-jump times is

$$M(t_{i+1}, -t_2) = \int_{t_i}^{+\infty} (t_{i+1}, -t_1) d(1 - t_i(t_{i+1})) = +\infty$$

which shows that the averages of inter-jump times is not useful. Also the median of t_{i+1} is $2 t_i$, independent of the time unit.

As $M(t_{i+1}, -t_i)^\alpha = t_i^\alpha \int_0^{+\infty} \xi^\alpha d(\frac{\xi}{1+\xi}) = t_i^\alpha \frac{\alpha \pi}{\sin \pi}$ if $\alpha < 1$ and nothing is gained because the effect of a change of time unit is cancelled; variance exists if $\alpha < 1/2$. Also we have

$$\text{Prob}\left\{\frac{t_{i+1}-t_i}{t_i} \leq \xi\right\} = \frac{\xi}{\xi+1}.$$

For other details see Deheuvels (1981, 1983) and references therein.

References

- Cramér, H., 1946. *Mathematical Methods of Statistics*, Princeton University Press, New Jersey.
- Dwass, M., 1964. Extremal processes. *Ann. Math. Statist.*, 35, 1718-1725.
- Dwass, M., 1966. Extremal processes, II, *Illinois J. Math.*, 10, 381-391.
- Dwass, M., 1973. Extremal processes, III, Northwestern Univ. Discussion Paper.
- Fisz, M., 1962. *Probability Theory and Mathematical Statistics*, Wiley, New York.
- Galambos, J., 1978. *The Asymptotic Theory of Extremes Order Statistics*, Wiley, New -York; (1987), 2nd ed., Krieger Publ.
- Gumbel, E. J., 1958. *Statistics of Extremes*, Columbia University Press, New York.
- Lamperti, J., 1964. On extreme order statistics. *Ann. Math. Statist.*, 35, 1726-1737.
- Leadbetter, M. R., Lindgren, G. and Rootzén, H., 1983. *Extremes and Related Properties in Random Sequences and Processes*, Springer-Verlag, Heidelberg.
- Mann, Henry B., 1953. *An Introduction to the Theory of Stochastic Process with Continuous Parameter*. Nat. Bur. Standards, Appl. Math. Series, (24), Washington.
- Newell, G. F., 1962. Asymptotic extreme value distribution for one-dimensional diffusion process. *J. Math. and Mech.*, 11, 481-496.
- Resnick, S. I. and Rubinovitch, M., 1973. The structure of extremal processes. *Adv. Appl. prob.*, 5, 287-307.
- Resnick, S. I., 1973. Extreme processes and record time values. *J. Appl. Prob.*, 10, 863-868.
- Resnick, S. I., 1975. Weak convergence to extremal processes. *Ann. Prob.*, 3, 951-1960.
- Selvey, S. D., 1975. *Statistical Inference*, Chapman and Hall, London.
- Tiago de Oliveira, J., 1972. Statistics for Gumbel and Fréchet distributions. *Structural Safety and Reliability*, A. M. Freudenthal ed., 91-105.
- Tiago de Oliveira, J., 1982. A definition of estimator efficiency in k-parameter case. *Ann. Inst. Statist. Math.*, 14,(3), A, 411-421, Tokyo.
- Tiago de Oliveira, J., 1962/63. Structure theory of bivariate extremes: extensions. *Estudos Mat., Estatist. Econometria*, VII, 165-194, Lisboa.
- Tiago de Oliveira, J., 1968. Extremal processes; definition and properties. *Publ. Inst. Statist. Univ. Paris*, XII, (2), 25-36.
