DOUBLY STOCHASTIC COMPOUND POISSON PROCESSES
IN EXTREME VALUE THEORY

H. Ferreira

Abstract: For some linear models, chain-dependent sequences and doubly stochastic
max-autoregressive processes, which do not satisfy the long range dependence condition
Δ(υ_n) from Hsing et al. ([7]), the sequence \{S_n\}_{n \geq 1}, of point processes of exceedances
of a real level υ_n by X_1, ..., X_n, n \geq 1, converges in distribution to a compound Poisson
process with stochastic intensity.

These examples illustrate the main result of this paper: for sequences \{X_n\}_{n \geq 1}
that conditional on a random variable X satisfy the usual dependence conditions in
the extreme value theory, we obtain the convergence of \{S_n\} to a point process whose
distribution is a mixture of distributions of compound Poisson processes. Such result
permits the identification of a class of sequences for which the extremal behaviour can
be described by mixtures of extreme value distributions.

1 – Introduction

Let \{u_n\}_{n \geq 1} be a sequence of real numbers and \{X_n\}_{n \geq 1} a sequence of random
variables such that, for some random variable X, every family of conditional
distributions given X,

\[ P_x = \left\{ P(X_{i_1} = x_{i_1}, ..., X_{i_n} = x_{i_n}) \mid n \geq 1, i_1 < \cdots < i_n \right\} \]  

(x in the support of X)
satisfies the dependence condition Δ(υ_n) introduced in Hsing et al. ([7]). That
is, if \( \{X_n^x\}_{n \geq 1} \) is a sequence with distribution determined by \( \mathcal{P}_x \) then
\[
\alpha_{n,\ell} = \alpha_{n,\ell}(x) = \sup \left\{ \left| P(A, B) - P(A)P(B) \right| : A \in \mathcal{B}_k^n(u_n), \ B \in \mathcal{B}_{k+\ell}^n(u_n), \ 1 \leq k \leq n - \ell \right\}, \quad 1 \leq \ell \leq n - 1,
\]
where \( \mathcal{B}_i^n(u_n) \) denotes the \( \sigma \)-field generated by \( \{X_i^x \leq u_n\}, \ i \leq s \leq j \), satisfies
\[
\alpha_{n,\ell_n} \rightarrow 0, \quad \text{for some sequence } \ell_n = o(n).
\]
From the results in Hsing et al. ([7]) it follows that, if \( \{X_n^x\}_{n \geq 1} \) is a stationary sequence and the sequence of point processes of exceedances
\[
S_n[X_n^x, u_n](\cdot) = \sum_{i=1}^{n} 1_{\{X_i^x > u_n\}} \delta_{u_n}(\cdot), \quad n \geq 1,
\]
converges in distribution to a point process \( S^x \), then \( S^x \) is a compound Poisson process \( S[\nu, \Pi] \) with Laplace transform
\[
L_{S^x}(f) = \exp \left( -\nu \int_{[0,1]} \left( 1 - \sum_{k=1}^{\infty} e^{-kf(y)} \Pi(k) \right) dy \right),
\]
where \( \nu = \nu(x) \) is a positive constant and \( \Pi = \Pi(\cdot) \) is a distribution for the multiplicities.

The most general results (Nandagopalan ([14]), Nandagopalan et al. ([15])) guarantee that for a non stationary sequence \( \{X_n^x\}_{n \geq 1} \) we can also find in the limit a compound process where Poisson events have a finite intensity measure \( \mu \) and the distributions of multiplicities \( \{\Pi_y\}_{y \in [0,1]} \) depend on the position \( y \) of the atoms, provided that some additional assumptions of equicontinuity and uniform asymptotic negligibility are satisfied.

Stronger results can be obtained if \( \{u_n\} \) is a sequence of normalized levels, \( \{u_n = u_n(\tau)\}_{n \geq 1} \), with \( \tau = \tau(x) \), for \( \{X_n^x\}_{n \geq 1} \), that is, if for each \( \alpha \in [0,1] \), it holds
\[
\sum_{i=1}^{[\alpha n]} P(X_i^x > u_n) \rightarrow a \alpha.
\]
If \( \{X_n^x\}_{n \geq 1} \) is a stationary sequence and if, for some \( \tau_0 > 0 \), \( \{S_n[X_n^x, u_n(\tau_0)]\}_{n \geq 1} \) converges in distribution, then for each \( \tau > 0 \), \( \{S_n[X_n^x, u_n(\tau)]\}_{n \geq 1} \) converges to \( S[\theta \tau, \Pi] \), with \( \theta = \frac{\tau_0}{\tau} \) and \( \Pi \) independent of \( \tau \).

This result of Hsing et al. ([7]) can be applied to some simple forms of non-stationarity like periodic sequences (Alpuim ([1]), Ferreira ([5])) and quasi-stationary sequences (Turkman ([16])).
A significant theory of point processes of rare events, under long range and local dependence conditions, is available in the recent literature and can be applied to obtain the asymptotic distribution $S^x$ of $S_n[X_n^x, u_n]$ (Hüsler, J. ([9]), Falk et al. ([3]), Hüsler, J. and Schmidt, M. ([10])). A simple application of the dominated convergence theorem enables us to conclude that $S_n[X_n^x, u_n]$ will converge to a point process whose distribution is a mixture of the distributions of the processes $S^x$. Such point process exists when, for each simple function $f$, the Laplace transform of $S^x$ on $f$, $\tilde{L}_S(f)$ is a measurable function (Kallenberg, [11]).

We begin in section 2 with a general result on the convergence of $S_n^g$ to a mixture. As a corollary, when $X_n$ conditional on realizations of $X$, is stationary, we obtain a sufficient condition for the convergence of $S_n[X_n^x, u_n]$ to a doubly stochastic compound Poisson process, i.e., a point process whose distribution is a mixture of the distributions of $S[\nu(x), \Pi]$, the intensity measure being regulated by $X$.

As applications we shall study the point process of exceedances of $u_n$ generated by the linear sequences $X_n = Y_n + X$ by some sequences whose finite distributions are determined from the values of a sequence of discrete random variables $X = \{J_n\}_{n \geq 1}$ and, finally, exceedances by the max-autoregressive sequences $X_n = \max(Y_n, Y_{n-1}, ..., Y_{n-\lambda})$.

The sequences to which we apply the results of this paper do not satisfy, in general, the condition $\Delta(u_n)$. However, on what concerns the local restrictions on rapid oscillations of $\{X_n\}$ (conditions $D'(u_n)$ from Leadbetter ([12]), $D''(u_n)$ from Leadbetter and Nandagopalan ([13]), $D(k)(u_n)$ from Chernick et al. ([2]), $\tilde{D}(k)(u_n)$ from Ferreira ([5])), they can be satisfied provided the analogous condition holds for each sequence $\{X_n^x\}_{n \geq 1}$.

We say that the condition $\tilde{D}(k)(u_n)$ holds, for a stationary sequence $\{X_n^x\}$ satisfying $\Delta(u_n)$, if $k$ is the minimum positive integer for which there exists a sequence of positive integers $\{k_n\}_{n \geq 1}$ satisfying

$$k_n \xrightarrow{n \to \infty} \infty, \quad k_n \ell_n/n \xrightarrow{n \to \infty} 0, \quad k_n a_n \ell_n \xrightarrow{n \to \infty} 0, \quad k_n P(X_1^x > u_n) \xrightarrow{n \to \infty} 0$$

and

$$(1.1) \quad s_n^{(k)} = n \sum_{2 \leq j_1 < ... < j_k \leq \lfloor n/k_n \rfloor - 1} P(X_1^x > u_n, \bigcap_{i=1}^k \{X_{j_i}^x \leq u_n < X_{j_i+1}^x\}) \xrightarrow{n \to \infty} 0.$$ 

When $k = 1$ we get the condition $D''(u_n)$ from Leadbetter and Nandagopalan ([13]).
The condition $D^{(k)}(u_n)$, from Chernick et al. ([2]), holds for a stationary sequence $\{X^n_x\}$ if

$$\lim_{n \to \infty} n^{[n/k_n]} \sum_{j=k+1}^{\lfloor n/k_n \rfloor} P\left(X^*_j > u_n, X^*_2 \leq u_n, \ldots, X^*_k \leq u_n, X^*_j > u_n\right) = 0. \tag{1.2}$$

These local dependence conditions enable us to obtain $\nu(x)$ and $\Pi(x)$, in the limiting compound Poisson process $S^x$, from certain limiting probabilities easy to compute, and also give criteria for the existence of the parameter $\theta = \theta(x)$, the extremal index of $\{X^n_x\}$.

By applying Fatou’s lemma, if (1.1) or (1.2) holds for each of the sequences $\{X^n_x\}$, then $\{X_n\}$ can satisfy the analogous convergence, which gives information about the episodes of high values of $\{X_n\}$.

2 – Main result and examples

**Proposition 2.1.** Let $\{X_n\}_{n \geq 1}$ be a sequence of random variables for which there exists a random variable $X$ such that, for each $x$ in the support of $X$, the sequence $\{X^n_x\}_{n \geq 1}$ with distribution determined by $P_x$ satisfies $S^n_x[X^n_x(u_n) \overset{d}{\longrightarrow} S^x]$. Suppose that, for each continuous, positive and with compact support function $f$, $\phi(x) = L_{S^x} f$ is a measurable function.

Then $\{S^n_x[X^n_x(u_n)]\}_{n \geq 1}$ converges, in distribution, to a point process whose distribution is a mixture of the distributions of $S^x$ regulated by the distribution of $X$.

**Proof.** It is sufficient to prove that, for any finite number of disjoint intervals $I_1, \ldots, I_k$ in $[0, 1]$ and non negative integers $s_1, \ldots, s_k$, it holds

$$P\left(\cap_{j=1}^k \left\{S^n_x[X^n_x(u_n)(I_j) = s_j]\right\}\right) \overset{n \to \infty}{\longrightarrow} \int P\left(\cap_{j=1}^k \left\{S^x(I_j) = s_j\right\}\right) dP_X(x). \tag{2.1}$$

Since

$$P\left(\cap_{j=1}^k \left\{S^n_x[X^n_x(u_n)(I_j) = s_j]\right\}\right) = \int P\left(\cap_{j=1}^k \left\{S^n_x[X^n_x(u_n)(I_j) = s_j] \mid X=x\right\}\right) dP_X(x)$$

$$= \int P\left(\cap_{j=1}^k \left\{S^n_x[X^n_x(u_n)(I_j) = s_j]\right\}\right) dP_X(x),$$

then, by using the dominated convergence theorem, we conclude (2.1).
If \( \{X^x_n\} \) is a stationary sequence and satisfies the condition \( \Delta(u_n) \) then \( S^x \) is a compound Poisson process \( S[\nu(x), \Pi^{(x)}] \). Furthermore, if we suppose that \( u_n = u_n(\tau(x)) \) for \( \{X^x_n\} \), i.e. \( \{u_n\}_{n \geq 1} \) is a sequence of normalized levels for \( \{X^x_n\}_{n \geq 1} \), then \( \nu(x) = \theta(x) \tau(x) \), with \( 0 \leq \theta(x) \leq 1 \).

In some examples we find \( \Pi^{(x)} \) and \( \theta(x) \) independent of \( x \). For instance, if for each \( x, \{X^x_n\} \) satisfies the condition \( D(u_n) \), in Leadbetter ([12]), that is

\[
\lim_{n \to \infty} \sup_{n \leq k} \sum_{j=2}^{[n/k]} P \left( X^x_j > u_n, X^x_k > u_n \right) \to 0,
\]

then \( \theta(x) = 1 \) and \( \Pi(x)(1) = 1 \). By applying the Proposition 2.1, we conclude that in these cases \( S_n[X^x_n, u_n] \) converges in distribution to a Cox process with stochastic intensity \( \tau(X) \).

The Proposition 2.1 is too general to give an insight into the limiting point process. The following application is a more useful result since in special cases of interest the limiting point process \( S^x \) is a compound Poisson process.

**Corollary 2.1.** Let \( \{X_n\} \) be a sequence of random variables for which there exists a random variable \( X \) such that, for each \( x \) in the support of \( X \), the sequence \( \{X^x_n\}_{n \geq 1} \) with distribution determined by \( \mathcal{P}_x \) satisfies \( S_n[X^x_n, u_n] \xrightarrow{d} S[\theta \tau(x), \Pi] \), where \( \tau(x) \) is a measurable function.

Then \( \{S_n[X^x_n, u_n]\}_{n \geq 1} \) converges, in distribution, to a doubly stochastic compound Poisson process with intensity \( \theta \tau(X) \) and multiplicity distribution \( \Pi \).

In the following we present two examples illustrating the above corollary. We shall maintain the notation \( X \) for the random variable chosen for conditioning.

We first give a limit distribution for the point process of exceedances generated by a linear model \( X_n = Y_n + X, \ n \geq 1 \), where \( \{Y_n\} \) is a periodic sequence independent of the variable \( X \).

The second example concerns a sequence \( \{X_n\} \) whose finite distributions are determined by a sequence of discrete random variables \( X = \{J_n\}_{n \geq 1} \).

The asymptotic distribution of \( S_n \) in these examples has already been obtained by specific methods (Ferreira ([4]), Turkman and Duarte ([17])). Our aim is to present one single methodology supported by the above results.

**Example 2.1.** Let \( T \) be a positive integer and \( Y_n, \ n \geq 1 \), independent exponential variables with parameter \( \lambda_n > 0 \) such that \( \lambda_n = \lambda_{n+T}, \ n \geq 1 \).

Then \( \{Y_n\} \) a \( T \)-periodic sequence, that is, for every choice of integers \( i_1 < \cdots < i_n, (Y_{i_1}, \ldots, Y_{i_n}) \) and \( (Y_{i_1+T}, \ldots, Y_{i_n+T}) \) are identically distributed.
If \( \lambda = \min_{1 \leq i \leq T} \lambda_i \) and \( C \) is the number of \( \lambda_i \)'s, \( i \leq T \), equal to \( \lambda \), then we check directly that \( v_n = -\frac{1}{T} \log \frac{AT}{n} \), \( n \geq 1 \), is a sequence of normalized levels for \( \{Y_n\}_{n \geq 1} \) with \( \tau = CA \) and for \( \{\max(Y_n, Y_{n+1})\}_{n \geq 1} \) with \( \tau = 2CA \):

\[
(2.2) \quad \lim_{n \to \infty} \frac{[na]}{T} \sum_{i=1}^{T} P(\max(Y_i, Y_{i+1}) > v_n) = \lim_{n \to \infty} \frac{na}{T} \sum_{i=1}^{T} \left[ 1 - P(Y_i \leq v_n, Y_{i+1} \leq v_n) \right] = \\
= \lim_{n \to \infty} \frac{na}{T} \sum_{i=1}^{T} \left[ \left( \frac{AT}{n} \right)^{\lambda_i / \lambda} + \left( \frac{AT}{n} \right)^{\lambda_{i+1} / \lambda} - \left( \frac{AT}{n} \right)^{(\lambda_i + \lambda_{i+1}) / \lambda} \right] = 2CA .
\]

Let us consider the \( T \)-periodic sequence

\[ X_n = \max(Y_n, Y_{n+1}) + X, \quad n \geq 1 , \]

where \( X \) is independent of \( \{Y_n\} \).

Then \( \{X_n^x\}_{n \geq 1} \) is a \( T \)-dependent and 2-dependent sequence. Applying (2.2) for \( u_n = -\frac{1}{T} \log \frac{1/2T}{n} \), we obtain

\[
\lim_{n \to \infty} \frac{[na]}{T} \sum_{i=1}^{T} P(X_i^x > u_n) = \lim_{n \to \infty} P(\max(Y_i, Y_{i+1}) + x > u_n) = aC e^{\lambda x} .
\]

Therefore, \( u_n = u_n(\tau) \) for \( \{X_n^x\} \) with \( \tau = C e^{\lambda x} \).

Simple calculations enables us to obtain

\[
n P(X_i^x \leq u_n < X_{i+1}^x) \xrightarrow{n \to \infty} \nu_i ,
\]

where \( \nu_i = \frac{1}{2} T e^{\lambda x} \) if \( \lambda_{i+2} = \lambda \) and \( \nu_i = 0 \) for the other cases;

\[
n T \sum_{i=1}^{T} P(X_i^x \leq u_n < X_{i+1}^x) \xrightarrow{n \to \infty} \nu = \frac{1}{2} \tau .
\]

Therefore, by applying the results in Ferreira ([5]), we conclude that, for each \( x \), it holds

\[
S_n[X_n^x, u_n] \xrightarrow{d}{n \to \infty} S[1/2 C e^{\lambda x}, \Pi], \quad \text{with} \quad \Pi(2) = 1 .
\]

and then \( \{S_n[X_n, u_n]\}_{n \geq 1} \) converges, in distribution, to a doubly stochastic compound Poisson process with intensity \( \frac{1}{2} C e^{\lambda x} \) and multiplicity distribution \( \Pi \). \( \diamond \)
A more interesting model corresponding to the situations where \( X \) and \( \{Y_n\} \) are not independent can be treated by using the ideas from the following example.

**Example 2.2.** Let \( X = \{J_n\}_{n \geq 1} \) be the following sequence of discrete random variables:

\[
P(J_n = k) = \delta_{n,k}, \quad \text{for } k = 1, \ldots, m, \quad \text{with } \sum_{k=1}^{m} \delta_{n,k} = 1, \quad n \geq 1.
\]

Consider a sequence \( \{X_n\}_{n \geq 1} \) whose finite distributions are determined by \( X = \{J_n\}_{n \geq 1} \) as follows:

\[
P\left(X_{i_1} \leq x_1, \ldots, X_{i_p} \leq x_p \mid \{J_n\}_{n \geq 1} = \{j_n\}_{n \geq 1}\right) = \prod_{s=1}^{p} P\left(X_{i_s} \leq x_s \mid J_{i_s} = j_{i_s}\right)
\]

\[
= \prod_{s=1}^{p} F_{j_{i_s}}(x_s),
\]

with non degenerated distributions \( F_1, \ldots, F_m \).

Therefore \( \{X_n^{(j_n)}\} \) is a sequence of independent and non-identically distributed variables with \( P(X_n^{(j_n)} \leq x) = F_{j_n}(x) \).

Suppose that, for each \( k \in \{1, \ldots, m\} \),

\[
\frac{1}{n} \sum_{i=1}^{n} 1_{(J_i = k)} \xrightarrow{a.s.} \frac{\text{a.s.}}{n \to \infty} N_k,
\]

where \( N_k \) is a positive random variable, and that

\[
n \left(1 - F_k(u_n)\right) \xrightarrow{n \to \infty} \tau_k > 0.
\]

Then, for almost every realization \( \{J_n\}_{n \geq 1} = \{j_n\}_{n \geq 1} \), if \( \eta_k = \lim_{n \to \infty} \frac{1}{n} \eta_{k,n} = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} 1_{(j_i = k)} \), for each \( k \in \{1, \ldots, m\} \), it holds

\[
\sum_{i=1}^{[na]} P\left(X_i^{(j_n)} > u_n\right) = \sum_{i=1}^{[na]} \frac{\tau_{j_i}}{n} + o(1) = \sum_{k=1}^{m} \frac{\eta_k \cdot [na]}{n} \tau_k + o(1) \quad \xrightarrow{n \to \infty} \quad a \sum_{k=1}^{m} \eta_k \tau_k = a \tau(\{j_n\}_{n \geq 1}).
\]

Therefore, \( u_n = u_n(\tau) \) for \( \{X_n^{(j_n)}\}_{n \geq 1} \).
Since \( \{X_n\} \) conditional on \( \{J_n\}_{n \geq 1} \) is a sequence of independent variables, then \( \sum_{X_n|J_n} u_n \) converges to an homogeneous Poisson process \( S[\tau] \), with intensity \( \tau \) (see Hüsler ([8])). Therefore, \( S_n[X_n, u_n] \) converges to a Cox process with stochastic intensity

\[
T = \tau(\{J_n\}) = \sum_{k=1}^{m} N_k \tau_k. 
\]

In the next example \( \theta(x) \) and \( \Pi(x) \) are not independent of \( x \) and we find the limiting multiplicities distributions with the help of results under the local dependence hypotheses.

**Example 2.3.** Let \( X \) be a random variable with values in \( \{1, 2\} \) and independent of a sequence \( \{Y_n\} \) of i.i.d. variables with continuous common d.f. Define

\[
X_n = \max(Y_n, Y_{n-X}, Y_{n-2}), \quad n \geq 1. 
\]

If \( u_n = u_n(\tau) \) for \( \{Y_n\} \) then \( u_n = u_n(3 \tau) \) for \( \{X_n^1\} \) and \( u_n = u_n(2 \tau) \) for \( \{X_n^2\} \).

The 2-dependent sequence \( \{X_n^1\} \) satisfies the condition \( D''(u_n) \) and, by applying the results in Leadbetter and Nandagopalan ([13]), we conclude that \( S[X_n^1, u_n] \xrightarrow{d} S[1/3, \tau, \Pi^{(1)}] \) with \( \Pi^{(1)}(3) = 1 \).

The 2-dependent sequence \( \{X_n^2\} \) does not satisfy the condition \( D''(u_n) \) but satisfies the condition \( D^{(3)}(u_n) \) from Chernick *et al.* ([2]) and the condition \( D^2(u_n) \) from Ferreira ([6]). It holds

\[
\lim_{n \to \infty} nP(X_1^2 > u_n) = 2 \tau, \\
\nu = \lim_{n \to \infty} nP(X_1^2 \leq u_n < X_2^2) = 2 \tau 
\]

and

\[
\beta = \lim_{n \to \infty} k_n \sum_{j=3}^{[n/k_n]-1} P(X_1^2 \leq u_n < X_2^2, \ X_j^2 \leq u_n < X_{j+1}^2) \\
= \lim_{n \to \infty} nP(X_1^2 \leq u_n < X_2^2, \ X_3^2 \leq u_n < X_4^2) = \tau, 
\]

for any sequence \( \{k_n\} \) of positive integers such that \( k_n \xrightarrow{n \to \infty} \infty \) and \( n/k_n \xrightarrow{n \to \infty} \infty \).
Then, applying directly the results in Ferreira ([6]) we obtain that \( \{X_n^2\} \) has extremal index \( \theta = \frac{\mu - \beta}{2r} = \frac{1}{2} \) and

\[
\Pi^{(2)}(\{1\}) = \lim_{n \to \infty} \frac{n}{\nu - \beta} P\left(X_1^2 \leq u_n, X_2^2 \leq u_n < X_3^2, X_4^2 \leq u_n, X_5^2 \leq u_n\right) = 0; \\
\Pi^{(2)}(\{2\}) = \lim_{n \to \infty} \frac{n}{\nu - \beta} P\left(X_1^2 \leq u_n < X_2^2, X_3^2 \leq u_n < X_4^2, X_5^2 \leq u_n\right) = 1.
\]

Therefore \( S_n[X_n, u_n] \) converges in distribution to a point process which distribution is a mixture of \( S[\frac{1}{3} \tau, \Pi^{(1)}] \) and \( S[\frac{1}{2} \tau, \Pi^{(2)}] \) directed by \( X \). 

ACKNOWLEDGEMENTS – I am grateful to the referee for his suggestions and corrections which helped in improving the final form of this paper.

REFERENCES


Helena Ferreira,
Universidade da Beira Interior,
R. Marquês d’Ávila e Bolama, 6200 Covilhã – PORTUGAL