Limit theorems for a recursive maximum process with location-dependent periodic intensity-parameter

Brice Franke*

December 5, 2008

Abstract

We investigate the recursive sequence $Z_n := \max\{Z_{n-1}, \lambda(Z_{n-1})X_n\}$ where X_n is a sequence of iid random variables with exponential distributions and λ is a periodic positive bounded measurable function. We prove that the Césaro mean of the sequence $\lambda(Z_n)$ converges toward the essential minimum of λ. Subsequently we apply this result and obtain a limit theorem for the distributions of the sequence Z_n. The resulting limit is a Gumbel distribution.

Keywords: Césaro convergence, maximum process, random recursion, extremal value

MSC: 60G70, 60F05

1 Introduction

The extremes of random recursions involving maxima of a sequence of iid random variables is a field having many practical applications (see for example Hooghiemstra and Keane (1985), Helland and Trygve (1976)). In this article we investigate a special type of recursions involving randomness. Let $\{X_n, n \in \mathbb{N}\}$ be a sequence of iid random variables having exponential distributions with parameter one and let $\lambda : \mathbb{R} \to \mathbb{R}$ be a positive bounded measurable and periodic function. We denote by $\lambda_{\inf}$ the essential infimum of λ on $\mathbb{R}^+$. Starting with $Z_0 = 0$ we define the following stochastic recursion equation:

$$Z_n := \max(Z_{n-1}, \lambda(Z_{n-1})X_n).$$

This is a Markov process on $\mathbb{R}^+$ with increasing paths. When the process is in location x at time n it waits until the first time when the sequence $\{\lambda(x)X_{n+k}; k \in \mathbb{N}\}$ exceeds the value x. If this happens at time $n+m$ then the process jumps from x to the new location $\lambda(x)X_{n+m}$

***Address:** Brice Franke, Fakultät für Mathematik, Ruhr-Universität Bochum, Universitätsstr. 150, 44780 Bochum, Germany; e-mail: Brice.Franke@rub.de

and waits there for the next jump.

In the next section we investigate the behavior of the string of observations $\{f(\lambda(Z_k)); k \in \mathbb{N}\}$, where f is a bounded measurable function which is right-continuous in λ_{inf}. For this we study the Césaro means

$$M_n := \frac{1}{n}\sum_{k=1}^{n} f(\lambda(Z_k)).$$

How does the sequence $\{M_n; n \in \mathbb{N}\}$ behave? First we observe that the larger the process $\{Z_k; k \in \mathbb{N}\}$ growths the longer it takes for its next jump to occur. Since the process goes to infinity this means that the process has larger and larger times of constancy between its jumps. Those times of constancy also depend on the value of λ at the location where the process waits for its next jump. At locations where λ is small we have to wait for a longer time-period until we see the next jump. Therefore, we expect that the process has larger times of constancy when the process stays at locations where the function λ is near its bottom λ_{inf}. This would imply that the Césaro mean has larger and larger proportions, where values of f near λ_{inf} are added. From this would result the convergence of the Césaro mean toward $f(\lambda_{\text{inf}})$. This is the statement of our first theorem. The arguments that we just developed are however not water proved. They do not take into account the possibility that after having accumulated a certain number of small values of λ the process makes a very large jump into a region with large values of λ. Since at far locations the hurdle that the process has to take is large, this also results in a long waiting time until the next overshoot. In that case the process might have to wait for a very long time before it jumps back into a region with small values of λ. Ruling out this possibility is one of the main difficulties in the proof of Theorem 1 (see Lemma 3).

In a second theorem we investigate for $a < b$ the behavior of the sequence

$$M_n(a,b) := \frac{1}{n}\sum_{k=[na]}^{[nb]} f(\lambda(Z_k)).$$

This case is more difficult to handle than Theorem 1, since it is not clear that during the time interval $\{[na], [na]+1, ..., [nb]\}$ the process jumps into the region where λ is small. In our second theorem we prove that $M_n(a,b)$ converges in probability toward $(b-a)f(\lambda_{\text{inf}})$.

In the final section we apply our main result in order to obtain a limit-law for the sequence of random variables $\{Z_n; n \in \mathbb{N}\}$. The resulting limit distribution will turn out to be a Gumbel distribution.

A number of questions seem to arise for further studies. It is probably possible to prove similar results to Theorem 1 for sequences $\{X_n; n \in \mathbb{N}\}$ with more general marginal distributions. The proof of such results however requires a much more technical approach, since we used the no-memory property of the exponential distribution at several places in our proof. It could be that the tail decay of such distributions has an effect on the limit behavior of the Césaro sum, since heavier tails might increase the chance that the process $\{Z_n; n \in \mathbb{N}\}$ makes large jumps after having accumulated a certain amount of small values of λ. The statement in Lemma 3 might be wrong in such situations. It seems that the periodicity of λ can be replaced by a condition which makes sure that at far locations the function λ has sufficiently many values which are close to $\lambda_{\inf}$. However, in such cases a proof requires considerably more effort, since we can not rely on a nice Markovian structure like in the proof of Lemma 2. Another question is whether the string of observations $\{f(\lambda(Z_k)); k \in \mathbb{N}\}$ satisfies a limit theorem or whether a large deviation result can be proved for this sequence.

2 A Césaro convergence theorem

In this section we prove the following theorem.

Theorem 1 *For every bounded measurable function $f : \mathbb{R}^+ \to \mathbb{R}^+$ which is right-continuous in $\lambda_{\inf}$ follows*

$$\frac{1}{n}\sum_{k=1}^{n} f(\lambda(Z_k)) \longrightarrow f(\lambda_{\inf}) \quad \mathbb{P} - \textit{almost surely as } n \to \infty.$$

Proof: For an arbitrary $\delta > 0$ we can find a $\epsilon > 0$ with the property

$$\sup_{z \in [\lambda_{\inf}, \lambda_{\inf}+2\epsilon]} |f(z) - f(\lambda_{\inf})| < \delta.$$

We define through induction the jump-times

$$\sigma_m := \inf\{l > \sigma_{m-1} : Z_l \neq Z_{l-1}\},\ \sigma_0 := 0.$$

Moreover, we define the sequences of random variables $S_m := Z_{\sigma_m};\ m \in \mathbb{N}$. The time that the process $\{Z_k; k \in \mathbb{N}\}$ spends after the m-th jump in S_m is given by

$$T_m(S_m) := \sigma_{m+1} - \sigma_m,\ m \in \mathbb{N}.$$

Finally we use induction to define the waiting times

$$\tau_0 = 0,\ \tau_k := \inf\{m > \tau_{k-1} : S_m \in \Lambda_\epsilon\},\ \text{where } \Lambda_\epsilon := \{x \in \mathbb{R} : \lambda(x) < \underline{\lambda}\} \text{ and } \underline{\lambda} := \lambda_{\inf} + \epsilon.$$

The following proposition gives upper bounds for $T_m(S_m)$, when $\lambda(S_m)$ is larger than $\overline{\lambda} := \lambda_{\inf} + 2\epsilon$. This will be crucial for the proof of Theorem 1:

Proposition 1 *The probability space $(\Omega, \mathcal{A}, \mathbb{P})$ can be enlarged to a probability space $(\tilde{\Omega}, \tilde{\mathcal{A}}, \tilde{\mathbb{P}})$ in a way such that there exists on $(\tilde{\Omega}, \tilde{\mathcal{A}}, \tilde{\mathbb{P}})$ a family of random variables $\{\overline{T}_m(x); m \in \mathbb{N}, x > 0\}$ satisfying the following properties:*
i) One has $T_m(S_m)(\omega) \leq \overline{T}_m(S_m)(\omega)$ for all $\omega \in \Omega$ with $\lambda(S_m(\omega)) \geq \overline{\lambda}$.
ii) The map $x \mapsto \overline{T}_m(x)(\omega)$ is increasing for all $m \in \mathbb{N}$ and all $\omega \in \tilde{\Omega}$.
iii) For all $x > 0$ the sequence $\{\overline{T}_m(x); m \in \mathbb{N}\}$ is iid with marginals having a geometric distribution with parameter $\overline{q}(x) = \exp\left(-\frac{1}{\overline{\lambda}}x\right)$.

iv) For all $y > 0$ and $m, m_1, ..., m_{n-1}, n \in \mathbb{N}$ one has

$$\begin{aligned} & \mathbb{P}\Big(T_{\tau_k}(S_{\tau_k}) = m, \tau_{k+1} - \tau_k = n, \overline{T}_{\tau_k+1}(y) = m_1, ..., \overline{T}_{\tau_k+n-1}(y) = m_{n-1}\Big| S_{\tau_k} = x\Big) \\ = \ & \mathbb{P}\Big(T_0(x) = m\Big)\mathbb{P}\Big(\tau_{k+1} - \tau_k = n\Big| S_{\tau_k} = x\Big)\mathbb{P}\Big(\overline{T}_1(y) = m_1, ..., \overline{T}_{n-1}(y) = m_{n-1}\Big), \end{aligned}$$

where $T_0(x)$ is a random variable having a geometric distribution with parameter

$$q_x := \exp\left(-\frac{1}{\lambda(x)}x\right).$$

Proof: We enlarge the probability space $(\Omega, \mathcal{A}, \mathbb{P})$ to a probability space $(\tilde{\Omega}, \tilde{\mathcal{A}}, \tilde{\mathbb{P}})$ in such a way such that there exists on $(\tilde{\Omega}, \tilde{\mathcal{A}}, \tilde{\mathbb{P}})$ a family of random variables $\{X_{m,n}; m, n \in \mathbb{N}\}$ with the following properties:
1. The family $\{X_{m,n}; m, n \in \mathbb{N}\}$ is iid with marginals having an exponential distribution with parameter one.
2. The two families $\{X_{m,n}; m, n \in \mathbb{N}\}$ and $\{X_n; n \in \mathbb{N}\}$ are independent.
We now define the auxiliary sequence

$$\tilde{X}_{m,n} := \begin{cases} X_{\sigma_m+n} & \text{for } n < \sigma_{m+1} - \sigma_m \\ X_{m,n} & \text{for } n \geq \sigma_{m+1} - \sigma_m \end{cases}$$

and the family of random variables

$$\overline{T}_m(x) := \inf\{n \in \mathbb{N} : \overline{\lambda}\tilde{X}_{m,n} > x\}.$$

We first prove that the two families $\{\tilde{X}_{m,n}; m, n \in \mathbb{N}\}$ and $\{X_{\sigma_m}; m \in \mathbb{N}\}$ are independent and that the family $\{\tilde{X}_{m,n}; m, n \in \mathbb{N}\}$ is iid with marginals which are exponentially distributed with parameter one.
For some array $\{x_{i,j}, Z_i, z_{m+1}; 1 \leq i \leq m, 1 \leq j \leq n\}$ of real numbers, we define the sets

$$\mathcal{U}_{m,n} := \Big\{\tilde{X}_{i,j} \geq x_{i,j}, X_{\sigma_i} \geq z_i, X_{\sigma_m} \geq z_m; 1 \leq i \leq m-1, 1 \leq j \leq n\Big\}$$

and

$$\mathcal{V}_{m,n} := \Big\{\tilde{X}_{m,j} \geq x_{m,j}, X_{\sigma_{m+1}} \geq z_{m+1}; 1 \leq j \leq n\Big\}.$$

We note that:
1. For all $k_1, ..., k_m \in \mathbb{N}$ the set $\{\sigma_1 = k_1, ..., \sigma_m = k_m\}$ is measurable with respect to the σ-algebra $\sigma\{X_1, ..., X_{k_m}\}$.
2. For all $k_1, ..., k_m \in \mathbb{N}$ the set $\mathcal{U}_{m,n} \cap \{\sigma_1 = k_1, ..., \sigma_m = k_m\}$ is measurable with respect to the σ-algebra $\sigma\big(\{X_1, ..., X_{k_m}\} \cup \{X_{i,k}; 1 \leq i \leq m-1, k \in \mathbb{N}\}\big)$.
3. For all $k_1, ..., k_m \in \mathbb{N}$ the set $\mathcal{V}_{m,n} \cap \{\sigma_1 = k_1, ..., \sigma_m = k_m\}$ is measurable with respect to the σ-algebra $\sigma\big(\{X_{k_1}, ..., X_{k_m}\} \cup \{X_{k_m+l}; l \in \mathbb{N}\} \cup \{X_{m,k}; k \in \mathbb{N}\}\big)$.

For $M \geq m$ it follows from those observations that conditioned on $X_{\sigma_1} = y_1, ..., X_{\sigma_M} = y_M$ the two events $\mathcal{U}_{m,n} \cap \{\sigma_1 = k_1, ..., \sigma_m = k_m\}$ and $\mathcal{V}_{m,n} \cap \{\sigma_1 = k_1, ..., \sigma_m = k_m\}$ are independent. We compute the probability of the event $\mathcal{V}_{m,n} \cap \{\sigma_1 = k_1, ..., \sigma_m = k_m\}$ under the condition $X_{\sigma_1} = y_1, ..., X_{\sigma_M} = y_M$. For this we define

$$s_m := \max\{s_{m-1}, \lambda(s_{m-1})y_m\},\ s_0 = 0 \text{ and } r_m = \frac{1}{\lambda(s_m)}s_m.$$

Under those assumptions we have for $l \leq n$ that

$$X_{\sigma_{m+1}} \geq z_{m+1},\ \tilde{X}_{m,j} \geq x_{m,j} \text{ for } 1 \leq j \leq n \text{ together with } \sigma_{m+1} - \sigma_m = l$$

is equivalent to

$$r_m \geq X_{k_m+i} \geq x_{m,j} \quad \text{for all } i \in \{1, ..., l-1\}, \quad X_{k_m+l} \geq \max\{r_m, z_{m+1}\},$$

and

$$X_{m,k_m+j} \geq x_{m,j} \quad \text{for all } j \in \{l, ..., n\}.$$

Hence it follows for $l \leq n$ that

$$\begin{aligned} & \mathbb{P}\Big(\mathcal{V}_{m,n} \cap \{\sigma_1 = k_1, ..., \sigma_m = k_m\} \cap \{\sigma_{m+1} - \sigma_m = l\} \Big| X_{\sigma_1} = y_1, ..., X_{\sigma_M} = y_M\Big) \\ = \; & \mathbb{P}\Big(r_m \geq X_i^{(\mathrm{aux})} \geq x_{m,i},\ 1 \leq i \leq l-1,\ X_j^{(\mathrm{aux})} \geq x_{m,j},\ l \leq j \leq n,\ X_0^{(\mathrm{aux})} \geq r_m \vee z_{m+1}\Big), \end{aligned}$$

where $\{X_i^{(\mathrm{aux})}; i \in \mathbb{N}_0\}$ is an auxiliary iid sequence of random variables having an exponential distribution with parameter one.
For $l \geq n+1$ we have that

$$X_{\sigma_{m+1}} \geq z_{m+1},\ \tilde{X}_{m,j} \geq x_{m,j} \text{ for } 1 \leq j \leq n \text{ together with } \sigma_{m+1} - \sigma_m = l$$

is equivalent to

$$r_m \geq X_{k_m+i} \geq x_{m,i} \text{ for all } i \in \{1, ..., n\},\ r_m \geq X_{k_m+j} \text{ for all } j \in \{n+1, ..., l-1\}$$

and

$$X_{k_m+l} \geq \max\{r_m, z_{m+1}\}.$$

Hence it follows for $l \geq n+1$ that

$$\begin{aligned} & \mathbb{P}\Big(\mathcal{V}_{m,n} \cap \{\sigma_1 = k_1, ..., \sigma_m = k_m\} \cap \{\sigma_{m+1} - \sigma_m = l\} \Big| X_{\sigma_1} = y_1, ..., X_{\sigma_M} = y_M\Big) \\ = \; & \mathbb{P}\Big(r_m \geq X_i^{(\mathrm{aux})} \geq x_{m,i},\ 1 \leq i \leq n,\ r_m \geq X_j^{(\mathrm{aux})},\ n+1 \leq j \leq l-1,\ X_0^{(\mathrm{aux})} \geq r_m \vee z_{m+1}\Big), \end{aligned}$$

where again $\{X_i^{(\mathrm{aux})}; i \in \mathbb{N}_0\}$ is an iid sequence of random variables having an exponential distribution with parameter one.

Summing up the conditional probabilities over $l \in \mathbb{N}$ yields

$$\begin{aligned}
&\mathbb{P}\Big(\mathcal{V}_{m,n} \cap \{\sigma_1 = k_1, ..., \sigma_m = k_m\}\Big| X_{\sigma_1} = y_1, ..., X_{\sigma_M} = y_M\Big) \\
= \ &\sum_{l=1}^{\infty} \mathbb{P}\Big(\mathcal{V}_{m,n} \cap \{\sigma_1 = k_1, ..., \sigma_m = k_m\} \cap \{\sigma_{m+1} - \sigma_m = l\}\Big| X_{\sigma_1} = y_1, ..., X_{\sigma_M} = y_M\Big) \\
= \ &\mathbb{P}\Big(X_j^{(\mathrm{aux})} \geq x_{m,j}, \ \forall j \in \{1, ..., n\}\Big)\mathbb{P}\Big(X_0^{(\mathrm{aux})} \geq r_m \vee z_{m+1}\Big) \\
= \ &\exp\Big(-\sum_{j=1}^{n} x_{m,j}\Big)\exp\Big(-r_m \vee z_{m+1}\Big).
\end{aligned}$$

We have that

$$\sigma_1 = k_1, ..., \sigma_m = k_m \quad \text{together with} \quad X_{\sigma_1} = y_1, ..., X_{\sigma_M} = y_M \quad \text{imply} \quad \frac{1}{\lambda(S_m)} S_m = r_m.$$

We note that $\frac{1}{\lambda(S_m)} S_m$ is measurable with respect to the σ-algebra $\sigma\{X_{\sigma_1}, ..., X_{\sigma_m}\}$. Together with the conditional independence discussed above, this yields

$$\begin{aligned}
&\mathbb{P}\Big(\mathcal{U}_{m,n} \cap \mathcal{V}_{m,n} \cap \{\sigma_1 = k_1, ..., \sigma_m = k_m\}\Big| X_{\sigma_1}, ..., X_{\sigma_M}\Big) \\
= \ &\mathbb{P}\Big(\mathcal{U}_{m,n} \cap \{\sigma_1 = k_1, ..., \sigma_m = k_m\}\Big| X_{\sigma_1}, ..., X_{\sigma_M}\Big) \\
&\quad \cdot \mathbb{P}\Big(\mathcal{V}_{m,n} \cap \{\sigma_1 = k_1, ..., \sigma_m = k_m\}\Big| X_{\sigma_1}, ..., X_{\sigma_M}\Big) \\
= \ &\mathbb{P}\Big(\mathcal{U}_{m,n} \cap \{\sigma_1 = k_1, ..., \sigma_m = k_m\}\Big| X_{\sigma_1}, ..., X_{\sigma_M}\Big) \\
&\cdot \ \exp\Big(-z_{m+1} \vee \frac{S_m}{\lambda(S_m)}\Big)\exp\Big(-\sum_{j=1}^{n} x_{m,j}\Big).
\end{aligned}$$

Summation over all possible values of $k_1, ..., k_m \in \mathbb{N}$ and an induction argument now yields

$$\begin{aligned}
&\mathbb{P}\Big(\mathcal{U}_{m,n} \cap \mathcal{V}_{m,n}\Big| X_{\sigma_1}, ..., X_{\sigma_M}\Big) \\
= \ &\mathbb{P}\Big(\mathcal{U}_{m,n}\Big| X_{\sigma_1}, ..., X_{\sigma_M}\Big)\exp\Big(-z_{m+1} \vee \frac{S_m}{\lambda(S_m)}\Big)\exp\Big(-\sum_{j=1}^{n} x_{m,j}\Big) \\
= \ &\exp\Big(-z_1\Big)\exp\Big(-z_2 \vee \frac{S_1}{\lambda(S_1)}\Big) \ ... \ \exp\Big(-z_{m+1} \vee \frac{S_m}{\lambda(S_m)}\Big)\exp\Big(-\sum_{i=1}^{m}\sum_{j=1}^{n} x_{i,j}\Big).
\end{aligned}$$

Taking expectation now implies

$$\begin{aligned}
&\mathbb{P}\Big(\tilde{X}_{i,j} \geq x_{i,j}, X_{\sigma_i} \geq z_i, X_{\sigma_m+1} \geq z_{m+1}; 1 \leq i \leq m, 1 \leq j \leq n\Big) \\
= \ &\mathbb{E}\Bigg[\exp\Big(-z_1\Big)\exp\Big(-z_2 \vee \frac{S_1}{\lambda(S_1)}\Big) \ ... \ \exp\Big(-z_{m+1} \vee \frac{S_m}{\lambda(S_m)}\Big)\Bigg]\exp\Big(-\sum_{i=1}^{m}\sum_{j=1}^{n} x_{i,j}\Big)
\end{aligned}$$

From this follows that the family $\{\tilde{X}_{m,n}; m, n \in \mathbb{N}\}$ is iid with marginals having an exponential distribution with parameter one. Moreover, this also proves the independence of the two families $\{\tilde{X}_{m,n}; m, n \in \mathbb{N}\}$ and $\{X_{\sigma_k}; k \in \mathbb{N}\}$.

i) The first claim follows from the fact that $\overline{\lambda} \leq \lambda(S_m)$ implies

$$\begin{aligned}
T_m(S_m) &= \sigma_{m+1} - \sigma_m \\
&= \inf\{l \in \mathbb{N} : \lambda(S_m) X_{\sigma_m + l} > S_m\} \\
&\leq \inf\{l \in \mathbb{N} : \lambda(S_m) \tilde{X}_{m,l} > S_m\} \\
&\leq \inf\{l \in \mathbb{N} : \overline{\lambda} \tilde{X}_{m,l} > S_m\} \\
&= \overline{T}_m(S_m).
\end{aligned}$$

ii) We see immediately from the definition of $\overline{T}_m(x)$ that the map $x \mapsto \overline{T}_m(x)(\omega)$ is increasing for all $m \in \mathbb{N}$ and all $\omega \in \tilde{\Omega}$.

iii) The fact that the family $\{\tilde{X}_{m,n}; m, n \in \mathbb{N}\}$ is iid with marginals which are exponentially distributed with parameter one shows that the sequence $\{\overline{T}_m(x); m \in \mathbb{N}\}$ is iid with marginals having a geometric distribution with parameter $\overline{q}(x)$.

iv) We just saw that the sequence of random variables $\{X_{\sigma_k}; k \in \mathbb{N}\}$ and the family of random variables $\{\tilde{X}_{m,n}; m, n \in \mathbb{N}\}$ are independent.

The event $\{\tau_{k+1} - \tau_k = n, \tau_k = l\}$ is measurable with respect to $\sigma\{S_{\tau_k}, X_{\sigma_{l+1}}, ..., X_{\sigma_{l+n}}\}$. It follows from the definitions that $\{\overline{T}_{\tau_k + i}(y) = m_i, \tau_k = l\}$ is measurable with respect to $\sigma\{\tilde{X}_{l+i,j}; j \in \mathbb{N}\}$. Moreover, the event $\{T_{\tau_k}(S_{\tau_k}) = m, \tau_k = l\}$ is measurable with respect to the σ-algebra $\sigma\{S_{\tau_k}, X_{\sigma_l + 1}, ..., X_{\sigma_l + m}\}$. From the construction of the family $\{\tilde{X}_{m,n}; m, n \in \mathbb{N}\}$ it follows that for $i \in \mathbb{N}$ the σ-algebras

$$\sigma\{\tilde{X}_{l+i,j}; j \in \mathbb{N}\} \quad \text{and} \quad \sigma\{X_{\sigma_l + 1}, ..., X_{\sigma_l + m}, X_{\sigma_{l+1}}, X_{\sigma_{l+2}}, ..., X_{\sigma_{l+n}}\} \ \text{ are independent.}$$

Thus it follows

$$\begin{aligned}
& \mathbb{P}\Big(T_{\tau_k}(S_{\tau_k}) = m, \tau_{k+1} - \tau_k = n, \overline{T}_{\tau_k + 1}(y) = m_1, ..., \overline{T}_{\tau_k + n - 1}(y) = m_{n-1} \Big| S_{\tau_k} = x\Big) \\
= & \sum_{l=1}^{\infty} \mathbb{P}\Big(T_{\tau_k}(S_{\tau_k}) = m, \tau_{k+1} - \tau_k = n, \overline{T}_{l+1}(y) = m_1, ..., \overline{T}_{l+n-1}(y) = m_{n-1}, \tau_k = l \Big| S_{\tau_k} = x\Big) \\
= & \sum_{l=1}^{\infty} \mathbb{P}\Big(T_{\tau_k}(S_{\tau_k}) = m, \tau_{k+1} - \tau_k = n, \tau_k = l \Big| S_{\tau_k} = x\Big) \\
& \qquad \cdot \mathbb{P}\Big(\overline{T}_{l+1}(y) = m_1, ..., \overline{T}_{l+n-1}(y) = m_{n-1}\Big) \\
= & \ \mathbb{P}\Big(T_{\tau_k}(S_{\tau_k}) = m, \tau_{k+1} - \tau_k = n \Big| S_{\tau_k} = x\Big) \mathbb{P}\Big(\overline{T}_1(y) = m_1, ..., \overline{T}_{n-1}(y) = m_{n-1}\Big).
\end{aligned}$$

In the last step we used the fact that the sequence $\{\overline{T}_k(y); k \in \mathbb{N}\}$ is iid and that the set $\{\overline{T}_{l+1}(y) = m_1, ..., \overline{T}_{l+n-1}(y) = m_{n-1}\}$ and the random variable S_l are independent.

Moreover, we have that

$$\begin{aligned}
& \mathbb{P}\Big(T_{\tau_k}(S_{\tau_k}) = m, \tau_{k+1} - \tau_k = n \Big| S_{\tau_k} = x\Big) \\
= & \sum_{l=1}^{\infty}\sum_{j=1}^{\infty} \mathbb{P}\Big(T_{\tau_k}(S_{\tau_k}) = m, \tau_{k+1} - \tau_k = n, \tau_k = l, \sigma_l = j \Big| S_{\tau_k} = x\Big) \\
= & \sum_{l=1}^{\infty}\sum_{j=1}^{\infty} \mathbb{P}\left(\{\tau_k = l, \sigma_l = j\} \cap \mathcal{M}_{j,m} \cap \mathcal{N}_{j,m.l} \Big| S_{\tau_k} = x\right)
\end{aligned}$$

with

$$\mathcal{M}_{j,m} := \Big\{\lambda(x)X_{j+1} \le x, ..., \lambda(x)X_{j+m-1} \le x, \lambda(x)X_{j+m} > x\Big\}$$

and

$$\mathcal{N}_{j,m,l} := \Big\{S_{l+1} = \lambda(x)X_{j+m} \notin \Lambda_\epsilon, S_{l+2} \notin \Lambda_\epsilon, ..., S_{l+n-1} \notin \Lambda_\epsilon, S_{l+n} \in \Lambda_\epsilon\Big\},$$

where $\Lambda_\epsilon := \{y \in \mathbb{R} : \lambda(y) \le \underline{\lambda}\}$. The sequence $\{Z_n; n \in \mathbb{N}\}$ is a Markov process and the random variables $\{\sigma_m; m \in \mathbb{N}\}$ are Markov times. Thus the sequence $\{S_m; m \in \mathbb{N}\}$ is a Markov process with transition probability densities

$$p(x,y) = \frac{1}{\lambda(x)} \exp\Big(- \frac{1}{\lambda(x)}(y-x)\Big) \mathbb{1}_{[x,\infty)}(y).$$

It follows from the Markov property that

$$\begin{aligned}
& \mathbb{P}\left(\{\tau_k = l, \sigma_l = j\} \cap \mathcal{M}_{j,m} \cap \mathcal{N}_{j,m.l} \Big| S_{\tau_k} = x\right) \\
= & \mathbb{E}\Bigg[\mathbb{1}_{\{\tau_k = l, \sigma_l = j\}} \mathbb{1}_{(-\infty,x]}(\lambda(x)X_{j+1}) \; ... \; \mathbb{1}_{(-\infty,x]}(\lambda(x)X_{j+m-1}) \mathbb{1}_{(x,\infty)}(\lambda(x)X_{j+m}) \\
& \cdot \mathbb{1}_{\Lambda_\epsilon^c}(\lambda(x)X_{j+m}) \int_{\Lambda_\epsilon^c} ... \int_{\Lambda_\epsilon^c} \int_{\Lambda_\epsilon} p(\lambda(x)X_{j+m}, y_2) \; ... \; p(y_{n-1}, y_n) dy_n ... dy_1 \Bigg| S_{\tau_k} = x\Bigg].
\end{aligned}$$

Since the events $\{\tau_k = l, \sigma_l = j, \lambda(x)X_{j+1} \le x, ..., \lambda(x)X_{j+m-1} \le x\}$ and $\{\lambda(x)X_{j+m} > x\}$ are independent, we have that

$$\begin{aligned}
& \mathbb{E}\Bigg[\mathbb{1}_{\{\tau_k = l, \sigma_l = j\}} \mathbb{1}_{(-\infty,x]}(\lambda(x)X_{j+1}) \; ... \; \mathbb{1}_{(-\infty,x]}(\lambda(x)X_{j+m-1}) \mathbb{1}_{(x,\infty)}(\lambda(x)X_{j+m}) \\
& \cdot \mathbb{1}_{\Lambda_\epsilon^c}(\lambda(x)X_{j+m}) \int_{\Lambda_\epsilon^c} ... \int_{\Lambda_\epsilon^c} \int_{\Lambda_\epsilon} p(\lambda(x)X_{j+m}, y_2) \; ... \; p(y_{n-1}, y_n) dy_n ... dy_1 \Bigg| \lambda(x)X_{j+m} > x\Bigg] \\
= & \mathbb{E}\Big[\mathbb{1}_{\{\tau_k = l, \sigma_l = j\}} \mathbb{1}_{(-\infty,x]}(\lambda(x)X_{j+1}) \; ... \; \mathbb{1}_{(-\infty,x]}(\lambda(x)X_{j+m-1})\Big] \mathbb{1}_{(x,\infty)}(\lambda(x)X_{j+m}) \\
& \mathbb{E}\Bigg[\mathbb{1}_{\Lambda_\epsilon^c}(\lambda(x)X_{j+m} \int_{\Lambda_\epsilon^c} ... \int_{\Lambda_\epsilon^c} \int_{\Lambda_\epsilon} p(\lambda(x)X_{j+m}, y_2) \; ... \; p(y_{n-1}, y_n) dy_n ... dy_1 \Bigg| \lambda(x)X_{j+m} > x\Bigg].
\end{aligned}$$

It follows from the fact that the exponential distribution has no memory that

$$\begin{aligned}
&\mathbb{E}\Big[1\!\!1_{\Lambda_\epsilon^c}(\lambda(x)X_{j+m}\int_{\Lambda_\epsilon^c}...\int_{\Lambda_\epsilon^c}\int_{\Lambda_\epsilon} p(\lambda(x)X_{j+m},y_2)\ ...\ p(y_{n-1},y_n)dy_n...dy_1\Big|\lambda(x)X_{j+m}>x\Big] \\
= &\int_{\Lambda_\epsilon^c}\int_{\Lambda_\epsilon^c}...\int_{\Lambda_\epsilon^c}\int_{\Lambda_\epsilon} p(x,y_1)p(y_1,y_2)\ ...\ p(y_{n-1},y_n)dy_n...dy_1.
\end{aligned}$$

This then yields

$$\begin{aligned}
&\mathbb{P}\Big(T_{\tau_k}(S_{\tau_k})=m,\tau_{k+1}-\tau_k=n\Big|S_{\tau_k}=x\Big) \\
= &\sum_{l=1}^\infty\sum_{j=1}^\infty \mathbb{P}\left(\{\tau_k=l,\sigma_l=j\}\cap\mathcal{M}_{j,m}\cap\mathcal{N}_{j,m.l}\Big|S_{\tau_k}=x\right) \\
= &\sum_{l=1}^\infty\sum_{j=1}^\infty \mathbb{E}\Big[1\!\!1_{\{\tau_k=l,\sigma_l=j\}}1\!\!1_{(-\infty,x]}(\lambda(x)X_{j+1})\ ...\ 1\!\!1_{(-\infty,x]}(\lambda(x)X_{j+m-1})\Big] \\
&\quad\cdot\mathbb{E}\Big[1\!\!1_{(-\infty,x]}(\lambda(x)X_{j+m})\Big]\int_{\Lambda_\epsilon^c}\int_{\Lambda_\epsilon^c}...\int_{\Lambda_\epsilon^c}\int_{\Lambda_\epsilon} p(x,y_1)p(y_1,y_2)\ ...\ p(y_{n-1},y_n)dy_n...dy_1 \\
= &(1-q_x)^{m-1}q_x\mathbb{P}\Big(\tau_{k+1}-\tau_k=n\Big|S_{\tau_k}=x\Big) \\
= &\mathbb{P}\Big(T_0(x)=m\Big)\mathbb{P}\Big(\tau_{k+1}-\tau_k=n\Big|S_{\tau_k}=x\Big).
\end{aligned}$$

This finishes the proof of part 4. □

In order to minimize abuse of notation we will use the notation $(\Omega,\mathcal{A},\mathbb{P})$ for the enlarged probability space $(\tilde{\Omega},\tilde{\mathcal{A}},\tilde{\mathbb{P}})$ in the remaining part of the manuscript.
We will later need the following lemma on the increments

$$Y_m := \frac{1}{\lambda(S_{m-1})}\big(S_m-S_{m-1}\big),\ m\in\mathbb{N}.$$

Lemma 1 *The family of random variables $\{Y_m;m\in\mathbb{N}\}$ is iid with marginals having an exponential distribution with parameter one.*

Proof: In this proof we will need the fact that the exponential distribution has no memory. We saw in the proof of Proposition 1 that $\{S_m;m\in\mathbb{N}\}$ is a Markov process with transition probability densities

$$p(x,y)=\frac{1}{\lambda(x)}\exp\Big(-\frac{1}{\lambda(x)}(y-x)\Big)1\!\!1_{[x,\infty)}(y).$$

It is not difficult to see that the σ-algebras $\sigma(Y_1, ..., Y_{m-1})$ and $\sigma(S_1, ..., S_{m-1})$ are equal. Then it follows from the Markov property that

$$\begin{aligned}
\mathbb{P}\Big(Y_m \geq y_m \Big| \sigma(Y_1, ..., Y_{m-1})\Big) &= \mathbb{P}\Big(\frac{1}{\lambda(S_{m-1})}\big(S_m - S_{m-1}\big) \geq y_m \Big| \sigma(S_1, ..., S_{m-1})\Big) \\
&= \mathbb{P}\Big(\frac{1}{\lambda(S_{m-1})}\big(S_m - S_{m-1}\big) \geq y_m \Big| \sigma(S_{m-1})\Big) \\
&= \mathbb{P}\Big(S_m \geq \lambda(S_{m-1})y_m + S_{m-1} \Big| \sigma(S_{m-1})\Big) \\
&= \mathbb{E}\Big[\exp\Big(-\frac{1}{\lambda(S_{m-1})}\big(\lambda(S_{m-1})y_m + S_{m-1} - S_{m-1}\big)\Big)\Big] \\
&= \exp\big(-y_m\big).
\end{aligned}$$

An induction argument proves that the family of random variables $\{Y_m; m \in \mathbb{N}\}$ is iid with marginals having an exponential distribution with parameter one. $\square$

In the following we denote by λ_{sup} the essential supremum of the function λ. The choice of $\epsilon > 0$ at the beginning of the proof of Theorem 1 implies the following inequalities

$$\left|\frac{1}{n}\sum_{k=1}^{n} f(\lambda(Z_k)) - f(\lambda_{\text{inf}})\right| \leq \delta\frac{1}{n}(G_n + D_n) + 2\sup_{z\in[\lambda_{\text{inf}},\lambda_{\text{sup}}]} |f(z)|\frac{1}{n}B_n$$

with

$$G_n := \sum_{k=1}^{n} \mathbb{1}_{(-\infty,\lambda_{\text{inf}}+\epsilon)}(\lambda(S_k))T_k(S_k)$$

$$D_n := \sum_{k=1}^{n} \mathbb{1}_{[\lambda_{\text{inf}}+\epsilon,\lambda_{\text{inf}}+2\epsilon)}(\lambda(S_k))T_k(S_k)$$

and

$$B_n := \sum_{k=1}^{n} \mathbb{1}_{[\lambda_{\text{inf}}+2\epsilon,\infty)}(\lambda(S_k))T_k(S_k).$$

In the remaining part of this section we will study the relative behavior of the two sequences $\{B_n; n \in \mathbb{N}\}$ and $\{G_n; n \in \mathbb{N}\}$ as $n \to \infty$. In order to do so we will first find some upper bound $\overline{B}_n$ for B_n and some lower bound $\underline{G}_n$ for G_n. Later we will use those bounds to see that G_n dominates CB_n for arbitrary large $C > 0$ as $n \to \infty$. This will imply that in the previous decomposition of the Césaro mean the part including the G_n-sequence dominates the part containing the B_n-sequence. This will finally prove Theorem 1.

If we define the random variable

$$\overline{B}_n^{(o)} := \sum_{k=1}^{n} \mathbb{I}_{[\lambda_{\inf}+2\epsilon,\infty)}(\lambda(S_k))\overline{T}_k(S_k)$$

it follows from part 1 in Proposition 1 that

$$B_n \leq \overline{B}_n^{(o)}.$$

We already defined the waiting times

$$\tau_0 = 0 \quad \text{and} \quad \tau_k := \inf\{m > \tau_{k-1} : \lambda(S_m) < \underline{\lambda}\}.$$

Then we define the random variables

$$\underline{G}_n := \sum_{k:\tau_k \leq n} T_{\tau_k}(S_{\tau_k})$$

and

$$\overline{B}_n := \sum_{k:\tau_k \leq n} \left(\sum_{\tau_{k-1}<m<\tau_k} \overline{T}_m(S_m) \right).$$

It then follows that

$$\underline{G}_n \leq G_n \quad \text{and} \quad B_n \leq \overline{B}_n^{(o)} \leq \overline{B}_n.$$

We will see later in Lemma 4 that for all $C > 0$ we have that

$$\mathbb{P}\left(C \sum_{\tau_k<m<\tau_{k+1}} \overline{T}_m(S_m) > T_{\tau_k}(S_{\tau_k}) \quad \text{infinitely often} \right) = 0.$$

This will imply that

$$\mathbb{P}\Big(C\frac{1}{n}\overline{B}_n > \frac{1}{n}\underline{G}_n \quad \text{infinitely often}\Big) = 0$$

and thus

$$\mathbb{P}\Big(C\frac{1}{n}B_n > \frac{1}{n}G_n \quad \text{infinitely often}\Big) = 0.$$

Applying this to the constant $C := \delta/\big(2\sup_{z\in[\lambda_{\inf},\lambda_{\sup}]}|f(z)|\big)$ yields that $\mathbb{P}$-almost surely there exists an $n_0 \in \mathbb{N}$ with

$$\left| \frac{1}{n}\sum_{k=1}^{n} f(\lambda(Z_k)) - f(\lambda_{\inf}) \right| \;\leq\; \delta\frac{1}{n}(G_n + D_n) + \delta\frac{1}{n}G_n \;\leq\; 2\delta \quad \text{for all } n \geq n_0.$$

Since $\delta > 0$ was arbitrary this finishes the proof of Theorem 1. □

We now turn to the proof of Lemma 4. We have to find arguments to make sure that between the times τ_{k-1} and τ_k the process does not visit a cite where the waiting time to the next jump exceeds the time that the process had to wait in $S_{\tau_{k-1}}$. This would happen if the process $\{S_m; m \in \mathbb{N}\}$ would jump over $\overline{\lambda} S_{\tau_k}/\underline{\lambda}$ during the time-interval $[\tau_k, \tau_{k+1}]$. We will prove in Lemma 3 that this happens only a finite number of times. To prove Lemma 3 we first need the following lemma on the growth of the sequence $\{S_{\tau_k}; k \in \mathbb{N}\}$.

Lemma 2 *One has*

$$\mathbb{P}\left(S_{\tau_k} < \sqrt{k} \quad \text{infinitely often}\right) = 0.$$

Proof: We define the map $\mathrm{pr} : [0,\infty) \to [0,1); x \mapsto x \bmod 1$. The function λ is periodic with period one, hence $\hat{S}_k := \mathrm{pr}(S_k)$ is a Markov process on the interval $[0,1)$. For $x, y \in [0,1)$ the transition densities for the Markov-process $\{\hat{S}_k; k \in \mathbb{N}\}$ are given by

$$\hat{p}(x,y) := \sum_{m \in \mathbb{N}} \frac{1}{\lambda(x)} \exp\left(-\frac{1}{\lambda(x)}(y+m-x)\right).$$

Therefore

$$\inf_{x,y \in [0,1)} \hat{p}(x,y) > 0$$

and it follows that there exists an invariant measure π on $[0,1)$ with the property

$$\frac{1}{k}\sum_{i=1}^{k} f(\hat{S}_i) \longrightarrow \int_{[0,1)} f(x)\pi(dx) \quad \mathbb{P} - \text{almost surely as } k \to \infty,$$

for all bounded measurable functions $f : [0,1) \to \mathbb{R}$ (see Doob (1953) p.220). This yields that for the set

$$\Lambda_\epsilon := \{x \in \mathbb{R} : \lambda(x) \le \lambda_{\inf} + \epsilon\}$$

one has as $k \to \infty$

$$\frac{1}{k}\inf\{j \in \mathbb{N}_0 : \tau_j \ge k\} = \frac{1}{k}\mathrm{card}\{i \in \{1,...,k\} : S_i \in \Lambda_\epsilon\} = \frac{1}{k}\sum_{i=1}^{k} \mathbb{1}_{\mathrm{pr}(\Lambda_\epsilon)}(\hat{S}_i) \longrightarrow \pi(\mathrm{pr}(\Lambda_\epsilon)).$$

As $k \to \infty$ we have $\tau_k \to \infty$ almost surely with respect to $\mathbb{P}$. This implies that $\mathbb{P}$-almost surely

$$\frac{k}{\tau_k} = \frac{1}{\tau_k}\inf\{j \in \mathbb{N}_0 : \tau_j \ge \tau_k\} \longrightarrow \pi(\mathrm{pr}(\Lambda_\epsilon)) \quad \text{as } k \to \infty$$

and consequently one has $\mathbb{P}$-almost surely $\tau_k/k \longrightarrow (\pi(\mathrm{pr}(\Lambda_\epsilon)))^{-1}$ as $k \to \infty$. Moreover, we have by induction that

$$S_n = S_{n-1} + \lambda(S_{n-1})Y_n \ge S_{n-1} + \lambda_{\inf} Y_n \ge \lambda_{\inf}\sum_{i=1}^{n} Y_i.$$

It then follows from the law of iterated logarithm applied to the iid sequence $\{Y_n; n \in \mathbb{N}\}$ that for all $\nu > 0$ there exists a $k_0 \in \mathbb{N}$ such that

$$S_{\tau_k} \geq \lambda_{\inf} \sum_{i=1}^{\tau_k} Y_i \geq (\lambda_{\inf} - \nu)\left(\tau_k - \sqrt{2\tau_k \log\log \tau_k}\right) \quad \text{for all } k \geq k_0.$$

Since $\tau_k \to \infty$ as $k \to \infty$, we can assume without loss of generality that

$$(\lambda_{\inf} - \nu)\left(\tau_k - \sqrt{2\tau_k \log\log \tau_k}\right) \geq \sqrt{\pi(\mathrm{pr}(\Lambda_\epsilon))}\sqrt{\tau_k} + 1 \quad \text{for all } k \geq k_0.$$

The asymptotic behavior of τ_k that we discussed above then yields that

$$\sqrt{\pi(\mathrm{pr}(\Lambda_\epsilon))}\sqrt{\tau_k} + 1 \geq \sqrt{k} \quad \text{for sufficiently large } k \in \mathbb{N}.$$

This finishes the proof. $\square$

We now are in position to prove Lemma 3. It turns out that $\frac{S_{\tau_{k+1}-1}}{\overline{\lambda}-\rho}$ exceeds $\frac{S_{\tau_k}}{\underline{\lambda}}$ only a finite number of times. The proof is based on a Borel Cantelli argument.

Lemma 3 *For all* $0 < \rho < \overline{\lambda} - \underline{\lambda}$ *one has*

$$\mathbb{P}\left(\frac{S_{\tau_{k+1}-1}}{\overline{\lambda}-\rho} > \frac{S_{\tau_k}}{\underline{\lambda}} \quad \text{infinitely often}\right) = 0.$$

Proof: To prove this result we use Borel Cantelli. We define $\lambda_{\sup}$ to be the essential supremum of the bounded measurable function λ. We note that

$$\begin{aligned}
& \mathbb{P}\left(\frac{S_{\tau_{k+1}-1}}{\overline{\lambda}-\rho} > \frac{S_{\tau_k}}{\underline{\lambda}},\ S_{\tau_k} \geq \sqrt{k}\right) \\
= \ & \mathbb{P}\left(S_{\tau_k} + \sum_{m=\tau_k+1}^{\tau_{k+1}-1} \lambda(S_{m-1})Y_m > \frac{\overline{\lambda}-\rho}{\underline{\lambda}} S_{\tau_k},\ S_{\tau_k} \geq \sqrt{k}\right) \\
\leq \ & \mathbb{P}\left(\lambda_{\sup} \sum_{m=\tau_k+1}^{\tau_{k+1}-1} Y_m > \left(\frac{\overline{\lambda}-\rho}{\underline{\lambda}} - 1\right) S_{\tau_k},\ S_{\tau_k} \geq \sqrt{k}\right) \\
\leq \ & \mathbb{P}\left(\sum_{m=\tau_k+1}^{\tau_{k+1}-1} Y_m > \frac{\sqrt{k}}{\lambda_{\sup}}\left(\frac{\overline{\lambda}-\rho}{\underline{\lambda}} - 1\right)\right) \\
\leq \ & \sum_{l=1}^{\infty} \mathbb{P}\left(\sum_{m=\tau_k+1}^{\tau_k+l-1} Y_m > \frac{\sqrt{k}}{\lambda_{\sup}}\left(\frac{\overline{\lambda}-\rho}{\underline{\lambda}} - 1\right),\ \tau_{k+1} - \tau_k = l\right).
\end{aligned}$$

By Cauchy Schwarz inequality we obtain

$$\begin{aligned}
&\sum_{l=1}^{\infty} \mathbb{P}\left(\sum_{m=\tau_k+1}^{\tau_k+l-1} Y_m > \frac{\sqrt{k}}{\lambda_{\sup}} \left(\frac{\overline{\lambda}-\rho}{\underline{\lambda}} - 1 \right),\ \tau_{k+1}-\tau_k = l \right) \\
\leq\ & \sum_{l=1}^{\infty} \sqrt{\mathbb{P}\left(\sum_{m=1}^{l-1} Y_m > \frac{1}{\lambda_{\sup}} \left(\frac{\overline{\lambda}-\rho}{\underline{\lambda}} - 1 \right) \sqrt{k} \right)} \sqrt{\mathbb{P}\big(\tau_{k+1}-\tau_k = l\big)} \\
\leq\ & \sum_{l=1}^{\infty} \exp\left(-\frac{1}{2l\lambda_{\sup}} \left(\frac{\overline{\lambda}-\rho}{\underline{\lambda}} - 1 \right) \sqrt{k} \right) \sqrt{\mathbb{P}\big(\tau_{k+1}-\tau_k = l\big)}.
\end{aligned}$$

As in the proof of Lemma 2 we define

$$\Lambda_\epsilon := \{x \in \mathbb{R} : \lambda(x) \leq \underline{\lambda}\}$$

and use the strong Markov property for the Markov process $\{S_n; n \in \mathbb{N}\}$ to see that

$$\begin{aligned}
\mathbb{P}\big(\tau_{k+1}-\tau_k = l\big) &\leq \mathbb{P}\left(S_{\tau_k+1} \notin \Lambda_\epsilon, ..., S_{\tau_k+l-1} \notin \Lambda_\epsilon\right) \\
&= \mathbb{E}\left[\int_{\Lambda_\epsilon^c} ... \int_{\Lambda_\epsilon^c} p(S_{\tau_k}, dx_1) \cdot ... \cdot p(x_{l-2}, dx_{l-1}) \right],
\end{aligned}$$

where

$$p(x, dy) := \frac{1}{\lambda(x)} \exp\left(-\frac{1}{\lambda(x)}(y-x) \right) \mathbb{I}_{[x,\infty)}(y) dy$$

is the transition kernel of the Markov-process $\{S_n; n \in \mathbb{N}\}$. Since the set Λ_ϵ and the function λ are periodic it follows that the function

$$x \mapsto \int_{\Lambda_\epsilon^c} p(x, dy)$$

is periodic and bounded from above by a constant $\overline{p} < 1$. It then follows that

$$\mathbb{P}\big(\tau_{k+1}-\tau_k = l\big) \leq \overline{p}^{l-1}.$$

Together with the previous computation this implies

$$\mathbb{P}\left(\frac{S_{\tau_{k+1}-1}}{\overline{\lambda}-\rho} > \frac{S_{\tau_k}}{\underline{\lambda}}, S_{\tau_k} \geq \sqrt{k} \right) \leq \sum_{l=1}^{\infty} \exp\left(-\frac{\sqrt{k}}{2l\lambda_{\sup}} \left(\frac{\overline{\lambda}-\rho}{\underline{\lambda}} - 1 \right) \right) \overline{p}^{(l-1)/2}.$$

We see that

$$\begin{aligned}
\sum_{k=1}^{\infty} \mathbb{P}\left(\frac{S_{\tau_{k+1}-1}}{\overline{\lambda}-\rho} > \frac{S_{\tau_k}}{\underline{\lambda}}, S_{\tau_k} \geq \sqrt{k} \right) &= \sum_{k=1}^{\infty} \sum_{l=1}^{\infty} \exp\left(-\frac{\sqrt{k}}{2l\lambda_{\sup}} \left(\frac{\overline{\lambda}-\rho}{\underline{\lambda}} - 1 \right) \right) \overline{p}^{(l-1)/2} \\
&= \sum_{l=1}^{\infty} \overline{p}^{(l-1)/2} \sum_{k=1}^{\infty} \exp\left(-\frac{\sqrt{k}}{2l\lambda_{\sup}} \left(\frac{\overline{\lambda}-\rho}{\underline{\lambda}} - 1 \right) \right).
\end{aligned}$$

We now introduce a new summation index $m^2 = k$. Then

$$m \leq \sqrt{k} < \sqrt{k+1} < ... < \sqrt{k+n(m)} \leq m+1$$

implies

$$m^2 \leq k < k+1 < ... < k+n(m) \leq (m+1)^2 \leq m^2+2m+2.$$

If we define

$$q(l) := \exp\left(-\frac{1}{2l\lambda_{\sup}}\left(\frac{\overline{\lambda}-\rho}{\underline{\lambda}}-1\right)\right),$$

then it follows from the above consideration that

$$\sum_{k=1}^{\infty} \exp\left(-\frac{\sqrt{k}}{2l\lambda_{\sup}}\left(\frac{\overline{\lambda}-\rho}{\underline{\lambda}}-1\right)\right) \leq 2\sum_{m=0}^{\infty}(m+1)(q(l))^m = 2\frac{d}{dq}\bigg|_{q=q(l)}\frac{1}{1-q} = \frac{2}{(1-q(l))^2}.$$

Further, one has

$$\frac{1}{(1-q(l))^2} = \left(1-\exp\left(-\frac{1}{2l\lambda_{\sup}}\left(\frac{\overline{\lambda}-\rho}{\underline{\lambda}}-1\right)\right)\right)^{-2} \sim \left(\frac{1}{2l\lambda_{\sup}}\left(\frac{\overline{\lambda}-\rho}{\underline{\lambda}}-1\right)\right)^{-2} = l^2C^2$$

with

$$C = \frac{1}{2\lambda_{\sup}}\left(\frac{\overline{\lambda}-\rho}{\underline{\lambda}}-1\right).$$

Since $\overline{p} \in (0,1)$ it follows that

$$\sum_{k=1}^{\infty} \mathbb{P}\left(\frac{S_{\tau_{k+1}-1}}{\overline{\lambda}-\rho} > \frac{S_{\tau_k}}{\underline{\lambda}}, S_{\tau_k} \geq \sqrt{k}\right) \leq \sum_{l=1}^{\infty} \overline{p}^{(l-1)/2}\frac{C^2}{(1-q(l))^2} < \infty.$$

The Borel Cantelli Lemma now implies

$$\mathbb{P}\left(\frac{S_{\tau_{k+1}-1}}{\overline{\lambda}-\rho} > \frac{S_{\tau_k}}{\underline{\lambda}} \;\text{ and }\; S_{\tau_k} \geq \sqrt{k} \quad \text{infinitely often}\right) = 0.$$

An application of Lemma 1 now finishes the proof. $\square$

Lemma 4 *One has for all $C > 0$*

$$\mathbb{P}\left(C\sum_{\tau_k<m<\tau_{k+1}} \overline{T}_m(S_m) > T_{\tau_k}(S_{\tau_k}) \quad \text{infinitely often}\right) = 0.$$

Proof: We can compute for all $x > 0$ that

$$\mathbb{P}\left(C \sum_{\tau_k < i < \tau_{k+1}} \overline{T}_i\left(\frac{\overline{\lambda} - \rho}{\underline{\lambda}} x\right) > T_{\tau_k}(S_{\tau_k}) \Big| S_{\tau_k} = x\right)$$

$$= \sum_{m=1}^{\infty}\sum_{l=1}^{\infty} \mathbb{P}\left(C \sum_{\tau_k < i < \tau_{k+1}} \overline{T}_i\left(\frac{\overline{\lambda} - \rho}{\underline{\lambda}} x\right) > l, T_{\tau_k}(S_{\tau_k}) = l, \tau_{k+1} - \tau_k = m \Big| S_{\tau_k} = x\right).$$

It follows from part 4 of Proposition 1 that this is equal to

$$\sum_{m=1}^{\infty}\sum_{l=1}^{\infty} \mathbb{P}\left(C \sum_{i=1}^{m-1} \overline{T}_i\left(\frac{\overline{\lambda} - \rho}{\underline{\lambda}} x\right) > l\right) \mathbb{P}\Big(T_0(x) = l\Big) \mathbb{P}\Big(\tau_{k+1} - \tau_k = m \Big| S_{\tau_k} = x\Big)$$

$$\leq \sum_{l=1}^{\infty}(1 - q_x)^{l-1} q_x \sum_{m=2}^{\infty} \mathbb{P}\left(\sum_{i=1}^{m-1} \overline{T}_i\left(\frac{\overline{\lambda} - \rho}{\underline{\lambda}} x\right) > [l/C]\right) \overline{p}^{m-1}.$$

In the last step we used the inequality $\mathbb{P}\big(\tau_{k+1} - \tau_k = m \big| S_{\tau_k} = x\big) \leq \overline{p}^{m-1}$, which can be found in the proof of Lemma 2 and the fact that the random variable $T_0(x)$ from part 4 in Proposition 1 has a geometric distribution with parameter

$$q_x := \exp\left(-\frac{1}{\lambda(x)} x\right).$$

We know from part 3 of Proposition 1 that for $y > 0$ the random variables $\overline{T}_1(y), ..., \overline{T}_{m-1}(y)$ are iid with marginals having a geometric distribution with parameter $\overline{q}(y) = \exp\Big(-\frac{1}{\overline{\lambda}} y\Big)$. This is due to the fact that the random variable $\overline{T}_1(y)$ are the waiting time for a first success in a row of iid Bernoulli experiments with success probability $\overline{q}(y)$. The sum of $m-1$ iid copies of $\overline{T}_1(y)$ is just the waiting time for the $m-1$-th success in a row of iid Bernoulli experiments with success probability $\overline{q}(y)$. This random variable has a negative binomial distribution, i.e.:

$$\mathbb{P}\left(\sum_{i=1}^{m-1} \overline{T}_i(y) = j\right) = \binom{j-1}{m-2} (\overline{q}(y))^{m-1} (1 - \overline{q}(y))^{j-m+1}.$$

If we define

$$\overline{q}_x := \overline{q}\left(\frac{\overline{\lambda} - \rho}{\underline{\lambda}} x\right)$$

then this fact and the inequality

$$\binom{j-1}{m-2} = \frac{(j-1)!}{(j-m+1)!\,(m-2)!} \leq \frac{j^{m-2}}{(m-2)!}$$

imply for large $x > 0$ that

$$\begin{aligned}
&\mathbb{P}\left(C \sum_{\tau_k < i < \tau_{k+1}} \overline{T}_i \left(\frac{\overline{\lambda}-\rho}{\underline{\lambda}} x \right) > T_{\tau_k}(S_{\tau_k}) \middle| S_{\tau_k} = x \right) \\
\leq\ & \sum_{l=1}^{\infty} (1-q_x)^{l-1} q_x \sum_{m=2}^{\infty} \overline{p}^{m-1} \sum_{j=[l/C]}^{\infty} \binom{j-1}{m-2} \overline{q}_x^{m-1} (1-\overline{q}_x)^{j-m+1} \\
\leq\ & \sum_{l=1}^{\infty} (1-q_x)^{l-1} q_x \sum_{m=2}^{\infty} \overline{p}^{m-1} \sum_{j=[l/C]}^{\infty} \frac{j^{m-2}}{(m-2)!} \overline{q}_x^{m-1} (1-\overline{q}_x)^{j-m+1} \\
=\ & \overline{p}\ q_x \overline{q}_x (1-\overline{q}_x)^{-1} \sum_{l=1}^{\infty} (1-q_x)^{l-1} \sum_{j=[l/C]}^{\infty} (1-\overline{q}_x)^j \exp\left(j \frac{\overline{q}_x \overline{p}}{(1-\overline{q}_x)} \right) \\
\leq\ & \frac{\overline{p}\ q_x \overline{q}_x (1-\overline{q}_x)^{-1}}{1-(1-\overline{q}_x)\exp\left(\frac{\overline{q}_x \overline{p}}{(1-\overline{q}_x)}\right)} \sum_{l=1}^{\infty} (1-q_x)^{l-1} (1-\overline{q}_x)^{[l/C]} \exp\left([l/C] \frac{\overline{q}_x \overline{p}}{(1-\overline{q}_x)} \right).
\end{aligned}$$

Here we used that for $x \to \infty$ we have $\overline{q}_x \to 0$ and thus for large $x > 0$

$$(1-\overline{q}_x)\exp\left(\frac{\overline{q}_x \overline{p}}{1-\overline{q}_x}\right) = (1-\overline{q}_x)\Big(1 + \frac{\overline{q}_x \overline{p}}{1-\overline{q}_x} + O(\overline{q}_x^2)\Big) = (1-\overline{q}_x) + \overline{q}_x \overline{p} + O(\overline{q}_x^2) < 1.$$

For $u \in (0,1)$ one has $u^{[l/C]} \leq u^{l/C-1} = u^{(l-1)/C + (1/C-1)}$. Thus we obtain that

$$\begin{aligned}
&\mathbb{P}\left(C \sum_{\tau_k < i < \tau_{k+1}} \overline{T}_i \left(\frac{\overline{\lambda}-\rho}{\underline{\lambda}} x \right) > T_{\tau_k}(S_{\tau_k}) \middle| S_{\tau_k} = x \right) \\
\leq\ & \frac{\overline{p}\ q_x \overline{q}_x (1-\overline{q}_x)^{1/C-2} \exp\left(\frac{\overline{q}_x \overline{p}}{1-\overline{q}_x}\right)}{\left(1-(1-\overline{q}_x)\exp\left(\frac{\overline{q}_x \overline{p}}{(1-\overline{q}_x)}\right)\right)\left(1-(1-q_x)(1-\overline{q}_x)^{1/C}\exp\left(\frac{\overline{q}_x \overline{p}}{C(1-\overline{q}_x)}\right)\right)}.
\end{aligned}$$

Taylor-approximation yields that $(1-\overline{q}_x)^{1/C} = 1 + O(\overline{q}_x)$ as $\overline{q}_x \to 0$. Further, we have that $\exp\left(\frac{\overline{q}_x \overline{p}}{(1-\overline{q}_x)}\right) = 1 + O(\overline{q}_x)$ as $\overline{q}_x \to 0$. It thus follows that

$$(1-q_x)(1-\overline{q}_x)^{1/C}\exp\left(\frac{\overline{q}_x \overline{p}}{C(1-\overline{q}_x)}\right) = (1-q_x)(1-O(\overline{q}_x))(1+O(\overline{q}_x)) = 1 - q_x + O(\overline{q}_x).$$

On the other hand we have $1-\exp\left(\frac{\overline{q}_x \overline{p}}{(1-\overline{q}_x)}\right) = O(\overline{q}_x)$ as $\overline{q}_x \to 0$ and it follows that

$$\left(1-(1-\overline{q}_x)\exp\left(\frac{\overline{q}_x \overline{p}}{(1-\overline{q}_x)}\right)\right) = O(\overline{q}_x) + \overline{q}_x O(1) = O(\overline{q}_x).$$

Finally we also have that

$$\overline{p}\; q_x \overline{q}_x (1-\overline{q}_x)^{1/C-2} \exp\left(\frac{\overline{q}_x \overline{p}}{1-\overline{q}_x}\right) = O(\overline{q}_x) q_x.$$

It then follows from those considerations that

$$\mathbb{P}\left(C \sum_{\tau_k < i < \tau_{k+1}} \overline{T}_i\left(\frac{\overline{\lambda}-\rho}{\underline{\lambda}} x\right) > T_{\tau_k}(S_{\tau_k}) \middle| S_{\tau_k} = x\right) \leq \frac{O(\overline{q}_x) q_x}{O(\overline{q}_x)\big(q_x + O(\overline{q}_x)\big)} = \frac{O(1) q_x}{\big(q_x + O(\overline{q}_x)\big)}.$$

This yields for $\overline{q}_x \to 0$ that

$$\mathbb{P}\left(C \sum_{\tau_k < i < \tau_{k+1}} \overline{T}_i\left(\frac{\overline{\lambda}-\rho}{\underline{\lambda}} x\right) > T_{\tau_k}(S_{\tau_k}) \middle| S_{\tau_k} = x\right) \leq \frac{O(1)}{\Big(1 + O(\overline{q}_x)/q_x\Big)}.$$

For $x \in \Lambda_\epsilon := \{z \in \mathbb{R} : \lambda(z) \leq \underline{\lambda}\}$ one has $q_x \leq \exp\left(-\frac{1}{\underline{\lambda}} x\right) =: \underline{q}_x$. Therefore, there exists in this situation a $K > 0$ with

$$\mathbb{P}\left(C \sum_{\tau_k < i < \tau_{k+1}} \overline{T}_i\left(\frac{\overline{\lambda}-\rho}{\underline{\lambda}} x\right) > T_{\tau_k}(S_{\tau_k}) \middle| S_{\tau_k} = x\right) \leq K \frac{\underline{q}_x}{\overline{q}_x} \leq K \exp\left(-\frac{\rho}{\overline{\lambda}\underline{\lambda}} x\right).$$

Since

$$\tau_k := \inf\{m > \tau_{k-1} : \lambda(S_m) < \underline{\lambda}\}$$

it follows that $S_{\tau_k} \in \Lambda_\epsilon$. Moreover, we saw in part 2 of Proposition 1 that $y \mapsto \overline{T}_l(y)$ is increasing. Those facts imply

$$\begin{aligned}
& \mathbb{P}\left(C \sum_{\tau_k < m < \tau_{k+1}} \overline{T}_m(S_m) > T_{\tau_k}(S_{\tau_k}),\ \frac{S_{\tau_{k+1}}}{\overline{\lambda}-\rho} \leq \frac{S_{\tau_k}}{\underline{\lambda}},\ S_{\tau_k} \geq \sqrt{k}\right) \\
\leq\ & \mathbb{P}\left(C \sum_{\tau_k < m < \tau_{k+1}} \overline{T}_m\left(\frac{\overline{\lambda}-\rho}{\underline{\lambda}} S_{\tau_k}\right) > T_{\tau_k}(S_{\tau_k}),\ S_{\tau_k} \geq \sqrt{k}\right) \\
=\ & \int_{\sqrt{k}}^{\infty} \mathbb{1}_{\Lambda_\epsilon}(x) \mathbb{P}\left(C \sum_{\tau_k < m < \tau_{k+1}} \overline{T}_m\left(\frac{\overline{\lambda}-\rho}{\underline{\lambda}} x\right) > T_{\tau_k}(S_{\tau_k}) \middle| S_{\tau_k} = x\right) \mathbb{P}\big(S_{\tau_k} \in dx\big) \\
\leq\ & K \int_{\sqrt{k}}^{\infty} \exp\left(-\frac{\rho}{\overline{\lambda}\underline{\lambda}} x\right) \mathbb{P}\big(S_{\tau_k} \in dx\big) \\
\leq\ & K \exp\left(-\frac{\rho}{\overline{\lambda}\underline{\lambda}} \sqrt{k}\right).
\end{aligned}$$

Since

$$\sum_{k=1}^{\infty} \exp\left(-\frac{\rho}{\overline{\lambda}\underline{\lambda}}\sqrt{k}\right) < \infty,$$

it follows from the Borel Cantelli lemma that

$$\mathbb{P}\left(C \sum_{\tau_k < m < \tau_{k+1}} \overline{T}_m(S_m) > T_{\tau_k}(S_{\tau_k}),\ \frac{S_{\tau_{k+1}}}{\overline{\lambda} - \rho} \leq \frac{S_{\tau_k}}{\underline{\lambda}},\ S_{\tau_k} \geq \sqrt{k} \ \text{ infinitely often} \right) = 0.$$

Applying Lemma 2 and Lemma 3 finishes the proof. □

3 A weak law for the Césaro convergence

We will need another version of the Césaro convergence theorem in the next section of this article. This version is different in two point from Theorem 1. The main difference is that we do not add up the full range $\{1, ..., n\}$ in the Césaro summation. We restrict the summation to $\{[an], [na] + 1, ..., [nb]\}$ where a and b are two positive real numbers. We are not able to prove almost sure convergence in this situation. So we have to restrain ourself to stochastic convergence, which is sufficient for the applications that we have in mind.

Theorem 2 *For all $a < b$ and every bounded measurable function $f : \mathbb{R}^+ \to \mathbb{R}^+$ which is right-continuous in λ_{inf} follows*

$$\frac{1}{n} \sum_{k=[na]}^{[nb]} f(\lambda(Z_k)) \longrightarrow (b-a) f(\lambda_{\text{inf}}) \quad \textit{in probability as } n \to \infty.$$

Proof: We use the same arguments as in the proof of Theorem 1 to see that for all $\delta > 0$ there exists a suitable $\epsilon > 0$ such that:

$$\left| \frac{1}{n} \sum_{k=[na]}^{[nb]} f(\lambda(Z_k)) - (b-a) f(\lambda_{\text{inf}}) \right| \quad \leq \quad \delta \frac{1}{n}(G_n + D_n) + 2 \sup_{z \in [\lambda_{\text{inf}}, \lambda_{\text{sup}}]} |f(z)| \frac{1}{n} B_n$$

with

$$G_n := \sum_{k=[na]}^{[nb]} \mathbb{1}_{(-\infty, \lambda_{\text{inf}} + \epsilon)}(\lambda(S_k)) T_k(S_k),$$

$$D_n := \sum_{k=[na]}^{[nb]} \mathbb{1}_{[\lambda_{\text{inf}} + \epsilon, \lambda_{\text{inf}} + 2\epsilon)}(\lambda(S_k)) T_k(S_k)$$

and

$$B_n := \sum_{k=[na]}^{[nb]} \mathbb{1}_{[\lambda_{\text{inf}} + 2\epsilon, \infty)}(\lambda(S_k)) T_k(S_k).$$

From this then follows that

$$\begin{aligned}
&\mathbb{P}\left(\left|\frac{1}{n}\sum_{k=[na]}^{[nb]} f(\lambda(Z_k)) - (b-a)f(\lambda_{\inf})\right| > 3\delta\right)\\
\leq\ & \mathbb{P}\left(\delta\frac{1}{n}(G_n + D_n) > \delta\right) + \mathbb{P}\left(2\sup_{z\in[\lambda_{\inf},\lambda_{\sup}]}|f(z)|\frac{1}{n}B_n > 2\delta\right)\\
\leq\ & 0 + \mathbb{P}\left(2\sup_{z\in[\lambda_{\inf},\lambda_{\sup}]}|f(z)|\frac{1}{n}\overline{B}_n > 2\delta\right)
\end{aligned}$$

since

$$\overline{B}_n := \sum_{k:[na]\leq\tau_k\leq[nb]}\left(\sum_{\tau_{k-1}<m<\tau_k}\overline{T}_m(S_m)\right) > B_n.$$

In order to finish the proof, it suffices to prove that for $C := \delta/(2\sup_{z\in[\lambda_{\inf},\lambda_{\sup}]}|f(z)|)$ one has

$$\mathbb{P}(C\overline{B}_n > \underline{G_n}) \longrightarrow 0 \text{ as } n\to\infty,$$

where

$$\underline{G}_n := \sum_{k:[na]\leq\tau_k\leq[nb]} T_{\tau_k}(S_{\tau_k}) \ \leq\ G_n.$$

For all $\rho > 0$ and all $n\in\mathbb{N}$ there exists a maximal natural number $M_n\in\mathbb{N}$ such that $\mathbb{P}(\tau_{M_n} < [na]) < \rho$. It follows that M_n goes to infinity when $n\to\infty$. Moreover, we have that

$$\begin{aligned}
&\sum_{k\geq M_n}\mathbb{P}\left(C\sum_{\tau_{k-1}<m<\tau_k}\overline{T}_m(S_m) > T_{\tau_k}(S_{\tau_k})\right)\\
\leq\ & \sum_{k\geq M_n}\mathbb{P}\left(C\sum_{\tau_{k-1}<m<\tau_k}\overline{T}_m(S_m) > T_{\tau_k}(S_{\tau_k}),\ \frac{S_{\tau_{k+1}}}{\overline{\lambda}-\rho}\leq\frac{S_{\tau_k}}{\underline{\lambda}},\ S_{\tau_k}\geq\sqrt{k}\right)\\
&+\sum_{k\geq M_n}\mathbb{P}\left(\frac{S_{\tau_{k+1}}}{\overline{\lambda}-\rho} > \frac{S_{\tau_k}}{\underline{\lambda}}\right) + \sum_{k\geq M_n}\mathbb{P}\left(S_{\tau_k}<\sqrt{k}\right).
\end{aligned}$$

We saw in the proof of Lemma 4 that the first series on the right side is finite. For the second one we have that

$$\sum_{k\geq M_n}\mathbb{P}\left(\frac{S_{\tau_{k+1}}}{\overline{\lambda}-\rho} > \frac{S_{\tau_k}}{\underline{\lambda}}\right) \leq \sum_{k\geq M_n}\mathbb{P}\left(\frac{S_{\tau_{k+1}}}{\overline{\lambda}-\rho} > \frac{S_{\tau_k}}{\underline{\lambda}}, S_{\tau_k}\geq\sqrt{k}\right) + \sum_{k\geq M_n}\mathbb{P}\left(S_{\tau_k}<\sqrt{k}\right).$$

Here again the first sum on the right side is finite as we saw in the proof of Lemma 3. We saw in the proof of Lemma 2 that $S_{\tau_k}\geq\lambda_{\inf}(Y_1+...+Y_k)$. This implies

$$\mathbb{P}\left(S_{\tau_k}\leq\sqrt{k}\right)\leq\mathbb{P}\left(\lambda_{\inf}(Y_1+...+Y_k)\leq\sqrt{k}\right) = \mathbb{P}\left(\frac{1}{\sqrt{k}}\sum_{i=1}^{k}(Y_i-\mathbb{E}[Y_1]) < \frac{1}{\lambda_{\inf}} - \sqrt{k}\mathbb{E}[Y_1]\right).$$

and it follows by the central limit theorem that

$$\sum_{k \geq M_n} \mathbb{P}\left(S_{\tau_k} < \sqrt{k}\right) < \infty.$$

Finally, it follows that there exists a large $n \in \mathbb{N}$ such that

$$\begin{aligned}
\mathbb{P}\Big(C\overline{B}_n > \underline{G}_n\Big) &\leq \mathbb{P}\Big(\tau_{M_n} < [na]\Big) + \mathbb{P}\Big(C\overline{B}_n > \underline{G}_n, \tau_{M_n} \geq [na]\Big) \\
&\leq \rho + \mathbb{P}\left(\bigcup_{k \geq M_n} \Big\{C \sum_{\tau_{k-1} < m < \tau_k} \overline{T}_m(S_m) > T_{\tau_k}(S_{\tau_k})\Big\}\right) \\
&\leq \rho + \sum_{k \geq M_n} \mathbb{P}\left(C \sum_{\tau_{k-1} < m < \tau_k} \overline{T}_m(S_m) > T_{\tau_k}(S_{\tau_k})\right) \\
&\leq 2\rho.
\end{aligned}$$

Since ρ was arbitrary, the desired stochastic convergence follows. □

4 Convergence of the extremal process

In this section we prove a limit-theorem for the sequence of processes

$$Z_t^{(n)} := \frac{1}{\lambda_{\inf}}\big(Z_{[nt]} - \log n\big).$$

The sequence has some dependence coming from its Markovian structure. The extremes of sequences which are related to an underlying Markov chain have been investigated in a number of research articles (see Resnick and Neuts (1970), Denzel and O'Brian (1975), Turkman and Walker (1983), Turkman and Oliveira (1992)). In those articles a discrete Markov chain influences the behavior of a sequence of random variables. This is also the case in our model, but the influence occurs in a different way.

The sequence of random variables $\{\lambda_{\inf} X_n; n \in \mathbb{N}\}$ is iid and has marginals with distribution function $F(x) = 1 - \mathrm{e}^{-x/\lambda_{\inf}}$. If $a_n := \lambda_{\inf}$ and $b_n := \log n$ it follows that

$$F^n(a_n x + b_n) = \left(1 - \mathrm{e}^{-(x + \log n)}\right)^n = \left(1 - \frac{\mathrm{e}^{-x}}{n}\right)^n \longrightarrow \exp\left(-\mathrm{e}^{-x}\right) = G(x) \quad \text{as } n \to \infty.$$

It follows from the previous computation that the exponential distribution is in the domain of attraction of the double-exponential distribution $G(x) := \exp(-\mathrm{e}^{-x}), x \in \mathbb{R}$; i.e.:

$$\mathbb{P}\big(\max(\lambda_{\inf} X_1, ..., \lambda_{\inf} X_n) \leq a_n x - b_n\big) \longrightarrow G(x).$$

The distribution $G(x)$ is also called the Gumbel distribution in the literature. It follows from Theorem 1 and Theorem 2 that the sequence $\{Z_n; n \in \mathbb{N}\}$ stays at locations with small λ-values most of the time. This fact gives a strong hint that the sequence $\{Z_n; n \in \mathbb{N}\}$ behaves like the sequence $\{\max(\lambda_{\inf} X_1, ..., \lambda_{\inf} X_n); n \in \mathbb{N}\}$ on large scales. Thus we expect that the distributions of the sequence $\{\frac{1}{a_n}(Z_n - b_n); n \in \mathbb{N}\}$ converge toward a Gumbel distribution as $n \to \infty$. We prove this result in this section. In order to do so we first prove the convergence of a suitable sequence of point processes which is associated to the processes $\{Z_t^{(n)}; t \geq 0\}$. The convergence of the sequence will then follow from an application of the continuous mapping theorem. This indirect approach to prove a Lamperti theorem can be found in Resnick (1987).

There exists a finest topology on $(-\infty, \infty]$ such that $x \mapsto \cot(x)$ is a homeomorphism from $(0, \pi/2]$ to $(-\infty, \infty]$. We now define a suitable point process on $\mathbb{R}^+ \times (-\infty, \infty]$:

$$N_n(dt, dz) := \sum_{k \in \mathbb{N}} \delta_{\left(k/n, \frac{1}{a_n}(Z_k - \log n)\right)}(dt, dz).$$

Theorem 3 *The point processes $\{N_n; n \in \mathbb{N}\}$ converge toward the Poisson point process N_* on $[0, \infty) \times \mathbb{R}$ with intensity measure*

$$\nu([0, t] \times (x, \infty]) := te^{-x}$$

Proof: Since the the random measures $\{N_n; n \in \mathbb{N}\}$ are simple point processes it is sufficient to find a basis $\mathcal{T}$ of relative compact open sets in $\mathbb{R}^+ \times (-\infty, \infty]$ which is closed under finite intersections and finite unions having the following properties (see Resnick (1987) p.157):

i) $\mathbb{P}(N_*(\partial F) = 0) = 1$ for all $F \in \mathcal{T}$;

ii) $\lim_{n \to \infty} \mathbb{P}(N_n(F) = 0) = \mathbb{P}(N_*(F) = 0)$ for all $F \in \mathcal{T}$;

iii) $\lim_{n \to \infty} \mathbb{E}[N_n(F)] = \mathbb{E}[N_*(F)]$ for all $F \in \mathcal{T}$.

We call a set F a figure if it is a finite disjoint union of open relatively compact rectangles from $\mathbb{R}^+ \times (-\infty, \infty]$. The set of figures $\mathcal{T}$ is obviously closed under finite unions and intersections.

Moreover condition (i) is certainly fulfilled since the limit-process is a Poisson point-process with absolute continuous intensity measure. Let K be an element from $\mathcal{T}$, then the family

$$N_n^{(K)}(t) := N_n\big(K \cap ((0, t] \times (-\infty, \infty])\big)$$

defines a point process on $\mathbb{R}^+$. The stochastic process $\{N_n^{(K)}(t); t > 0\}$ has a canonical filtration $\mathcal{F}_t^{(K,n)} := \sigma\big(N_n^{(K)}(s); s \leq t\big)$. Let $\{A_n^{(K)}(t); t \geq 0\}$ be the compensator associated to the process $N_n^{(K)}$ and the filtration $\mathcal{F}^{(K,n)}$ (see Daley and Vere-Jones (1988) p.514). We want to obtain an explicit representation for the compensator $A^{(K)}$. The figure K has the following decomposition

$$K = \bigcup_{i=1}^{M} \Big(T_i \times Q_i\Big),$$

where $T_i \subset \mathbb{R}^+$ and $Q_i \subset (-\infty, \infty]$ are open relatively compact intervals. The compensator for $N_n^{(K)}$ has the following expression (see Jacod and Shiryaev (2003) p.94)

$$\begin{aligned} A_n^{(K)}(t) &= \sum_{i=1}^{M} \sum_{j/n \in T_i \cap [0,t]} \frac{\lambda_{\inf}}{\lambda(Z_{j-1})} \int_{Q_i} \exp\left(-\frac{\lambda_{\inf}}{\lambda(Z_{j-1})} y - \log n\right) dy \\ &= \frac{1}{n} \sum_{i=1}^{M} \sum_{j/n \in T_i \cap [0,t]} \frac{\lambda_{\inf}}{\lambda(Z_{j-1})} \int_{Q_i} \exp\left(-\frac{\lambda_{\inf}}{\lambda(Z_{j-1})} y\right) dy. \end{aligned}$$

By Theorem 2 this converges for every $t > 0$ in probability toward

$$A^{(K)}(t) := \sum_{i=1}^{M} |T_i \cap [0,t]| \int_{Q_i} \exp(-y)\, dy,$$

where we denote by $|I|$ the lenght of an interval $I \subset \mathbb{R}$. The last expression is just the compensator of the point-process

$$N_*^{(K)}(t) := N_*\big(K \cap ((0,t] \times (-\infty, \infty])\big).$$

The distributional convergence of the processes $N_n^{(K)}$ toward $N^{(K)}$ implies the distributional convergence of $N_n^{(K)}(\theta)$ toward $N^{(K)}(\theta)$ for $\theta := \sup_i \sup T_i$. This proves condition (ii), since

$$\lim_{n\to\infty} \mathbb{P}(N_n(K) = 0) = \lim_{n\to\infty} \mathbb{P}(N_n^{(K)}(\theta) = 0) = \mathbb{P}(N_*^{(K)}(\theta) = 0) = \mathbb{P}(N_*(K) = 0).$$

In order to prove condition (iii) we note that the fact that $A_n^{(K)}$ and $A_*^{(K)}$ are compensators yields

$$\mathbb{E}\left[N_n(K)\right] = \mathbb{E}\left[N_n^{(K)}(\theta)\right] = \mathbb{E}\left[A_n^{(K)}(\theta)\right] \quad \text{and} \quad \mathbb{E}\left[N_*(K)\right] = \mathbb{E}\left[N_n^{(K)}(\theta)\right] = \mathbb{E}\left[A_*^{(K)}(\theta)\right].$$

Condition (iii) then follows by Theorem 2. $\square$

To the distribution function G we associate an extreme-value process having finite dimensional distributions defined as follows

$$G_{t_1,\dots,t_k}(x_1, \dots, x_k) := G^{t_1}\left(\bigwedge_{i=1}^{k} x_i\right) G^{t_2 - t_1}\left(\bigwedge_{i=2}^{k} x_i\right) \cdot \dots \cdot G^{t_k - t_{k-1}}(x_k).$$

The resulting stochastic process $\{Z(t), t > 0\}$ is a Markov-process with non-decreasing paths. A version of this process exists in $D(0, \infty)$ (see Resnick (1987)). We now define the associated maximum processes

$$Z_t^{(n)} := \frac{1}{a_n}\Big(Z_{[nt]} - \log n\Big).$$

Corollary 1 *The processes $\{Z^{(n)}; n \in \mathbb{N}\}$ converge toward the extreme-value process associated to the Gumbel-distribution*

$$G(x) = \exp\left(-\mathrm{e}^{-x}\right).$$

Proof: For the proof we define the map

$$\mathfrak{F} : Z_p([0,\infty) \times (-\infty,\infty]) \to D(0,\infty); \mu = \sum_k \delta_{t_k,j_k} \mapsto \left(t \mapsto \bigvee_{0<t_k\le t} j_k \right).$$

It follows that $\mathfrak{F}(N_n) = Z^{(n)}$ (see Resnick (1987) p.209) and that $\mathfrak{F}$ is continuous almost everywhere with respect to the distribution of N_*. The continuous mapping theorem and the Theorem 2 then implies that $Z^{(n)}$ converges in distribution toward $Z = \mathfrak{F}(N)$. □

Corollary 2 *The sequence $\mathbb{P}\left(\frac{1}{a_n}\left(Z_n - \log n\right) > x\right)$ converges toward the Gumbel-distribution*

$$G(x) = \exp\left(-\mathrm{e}^{-x}\right).$$

Proof: This follows directly from the previous corollary. □

Acknowledgements: The author wants to thank Herold Dehling and Sheu Shuenn-Jyi for some interesting discussions during the preparation of this manuscript.

5 References

Daley, D.J.; Vere-Jones, D.: *An introduction to the theory of point processes*, Springer Series in Statistics. Springer-Verlag, New York, (1988).

Denzel G.E., O'Brien G.L.: Limit theorems for extreme values of chain dependent processes, *The Annals of Probability*, **3**, 773-779, (1975).

Doob J.L.: *Stochastic Processes*, Willey, New-York, (1953).

Helland I.s., Trygve S.N.: On a general random exchange model, *Journal of Applied Probability*, **13**, 781-790, (1976).

Hooghiemstra G., Keane M.: Calculation of the equilibrium distribution for a solar energy storage model, *Journal of Applied Probability*, **22**, 852-864, (1985).

Jacod, J.; Shiryaev, A.N.: *Limit theorems for stochastic processes*, Second edition, Grundlehren der Mathematischen Wissenschaften, **288**, Springer-Verlag, Berlin, (2003).

Resnick, S.I.: *Extreme Values, Regular Variation, and Point Processes*, Applied Probability, **4**, Springer-Verlag, New-York (1987).

Resnick S.I., Neuts M.F.: Limit laws for maxima of a sequence of random variables defined on a Markov chain, *Advances in Applied Probability*, **4**, 285-295, (1970).

Turkman K.F., Oliveira M.F.: Limit laws for the maxima of chain dependent sequences with positive extremal index, *Journal of Applied Probability*, **29**, 222-227, (1992).

Turkman K.F., Walker A.M.: Limit laws for the maxima of a class of quasi-stationary sequences, *Journal of Applied Probability*, **20**, 814-821, (1983).