

EVOLUTIONARY PARETO DISTRIBUTIONS

Dominique PICARD^a, Karine TRIBOULEY^b

^a *Laboratoire de Probabilités et Modèles Aléatoires, CNRS-UMR 7599,
Université Paris VI et Université Paris VII, 4, Place Jussieu, Tour 56 3eme étage,
F-75252 Paris Cedex 05, France*

^b *URA CNRS 743, Université Paris XI, Bâtiment 425, 91405 Orsay Cedex, France*

Received 29 May 2001, revised 22 January 2002

ABSTRACT. – We consider here i.i.d. variables which are distributed according to a Pareto $\mathcal{P}(\alpha)$ up to some point x_1 and a Pareto $\mathcal{P}(\beta)$ (with a different parameter) after this point. This model constitutes an approximation for estimating extreme tail probabilities, especially for financial or insurance data. We estimate the parameters by maximizing the likelihood of the sample, and investigate the rates of convergence and the asymptotic laws. We find here a problem which is very close to the change point question from the point of view of limits of experiments. Especially, the rates of convergence and the limiting law obtained here are the same as in a change point framework. Simulations are giving an illustration of the quality of the procedure.

© 2002 Éditions scientifiques et médicales Elsevier SAS

Keywords: Change point; Schauder basis; Likelihood process; Pareto distribution

RÉSUMÉ. – Nous considérons un n -échantillon de variables aléatoires réelles qui suivent une loi de Pareto $\mathcal{P}(\alpha)$ jusqu'à un point x_1 puis une Pareto $\mathcal{P}(\beta)$ après ce point. Ce modèle est une approximation des modèles utilisés pour estimer des probabilités d'extrêmes en particulier en finances.

Nous considérons les estimateurs des différents paramètres qui maximisent la vraisemblance et déterminons leur lois asymptotiques. Nous retrouvons un problème proche du problème de rupture de modèle du point de vue des expériences limites. En particulier, les vitesses de convergence des estimateurs ainsi que les lois limites obtenues sont les mêmes. Nous donnons des simulations pour montrer la qualité de cette procédure.

© 2002 Éditions scientifiques et médicales Elsevier SAS

1. Introduction

Our aim in this paper is to analyze the likelihood process of n independent identically variables which common law has the following distribution function

$$F_{x_0, x_1, \alpha, \beta}(x) = \left(1 - \left(\frac{x_0}{x}\right)^\alpha\right) 1_{\{x_0 \leq x \leq x_1\}} + \left(1 - \left(\frac{x_0}{x_1}\right)^\alpha \left(\frac{x_1}{x}\right)^\beta\right) 1_{\{x_1 \leq x\}} \quad (1)$$

for $0 < x_0 < x_1$ and $\alpha, \beta \geq 1$. We are interested in estimating the parameters α, β, x_0, x_1 . As can be seen, this distribution clearly presents two regimes, each of them is of Pareto

type. The parameter x_1 appears as particularly challenging to estimate since it is the point of change between the two different regimes. We show that this parameter actually behaves like a change point parameter, in the sense of experiments. More precisely, we show that the likelihood process converges in the same sense and to the same limiting process as in the case of a change point problem.

Hence, we find results which are strongly connected to Deshayes, Picard [3,4]. The technic of proof also follows rather closely the previous work. However, for the proof of the tightness, we take advantage of the now classical technology of the wavelets (here the Ciesielski basis), and the associated spaces.

The work cited above was part of a thesis conducted with Didier Dacunha Castelle as advisor. It was also very much inspired by the work of Lucien LeCam (see [8]) on limits of experiments and of Ildar Ibragimov and Rafael Hasminskii (see [7]) on the likelihood process.

However, besides this homage, we had specific motivations. A very important problem in finance or insurance consists in estimating values at risk (VaR), or in other words quantiles for very small (or very big) probabilities (see for instance the impressive book of Embrecht, Klüppelberg and Mikosch [5]). As we are concerned with tails, the theorems of attraction for the maximum of n variables are generally applying, and under assumptions which are generally fulfilled in the context of insurance, the problem roughly consists in estimating the tail index of a Pareto distribution. However, this estimation has to be performed on the observations where the attraction is supposed to be true, i.e. the ‘tail’ observations. Immediately, a very delicate problem occurs in theory and even more in practice: how to choose the number of such observations? There obviously appears a trade off between a bias phenomenon (the attraction is reasonable only for the very last observations) and variance (the larger the number of taken observations, the smaller is the variance). Various possibilities have been proposed, let us just cite as a reference the fundamental paper of de Haan, Peng [2]. We adopt here a rather different point of view, by trying to find this optimal number of observations as the point of change between two regimes, where the last one obviously is a Pareto. We do not give here a complete resolution of this aspect since we are only considering two Pareto regimes. The next step consisting in performing the procedure with a much broader assumption on the first regime, will be done in a following paper.

2. Study of the likelihood ratio process

As in the context of a change point model, one can prove here that the interesting situation corresponds to assuming that $d_n = \beta - \alpha$ (hence in fact $\beta_n - \alpha_n$) is tending to zero when n is tending to infinity. Let u, v, w, t be some constants and let us consider the sequence of processes:

$$\Gamma_n(u, v, w, t) = \frac{L_n(x_0 + \frac{w}{n}, x_1 + \frac{t}{nd_n^2}, \alpha_n + \frac{u}{\sqrt{n}}, \beta_n + \frac{v}{\sqrt{n}})}{L_n(x_0, x_1, \alpha_n, \beta_n)}$$

where $L_n(x_0, x_1, \alpha, \beta)$ is the likelihood of the sample X_1, \dots, X_n (i.i.d. variables of distribution given by (1)). We obtain the following result:

THEOREM 1. – *Let $0 < x_0 < x_1$ be real numbers. Let us suppose that the real parameters α_n, β_n are larger than 1 and are depending on n in such a way that $\lim_{n \rightarrow +\infty} \alpha_n = \lim_{n \rightarrow +\infty} \beta_n = \alpha^*$ with $d_n = \beta_n - \alpha_n$ tending to 0 and nd_n^4 tending to infinity.*

Then, the sequence of processes $\{\Gamma_n(u, v, w, t), (u, v, w, t) \in \mathbb{R}^4\}$ weakly converges under $\mathbb{P}_{x_0, x_1, \alpha_n, \beta_n}^n$ to the process $\{\Gamma(u, v, w, t), (u, v, w, t) \in \mathbb{R}^4\}$ with the topology of uniform convergence on compact sets, where Γ is defined by:

$$\log \Gamma(u, v, w, t) = \left[\left(u\xi_1 - \frac{1}{2}u^2\sigma_1^2 \right) + \left(v\xi_2 - \frac{1}{2}v^2\sigma_2^2 \right) + c \left(W(t) - \frac{|t|}{2} \right) + w \frac{\alpha^*}{x_0} \right] 1_{\{E > w\}}$$

where

- ξ_1 is a gaussian vector of mean 0 and of variance $\sigma_1^2 = \frac{1 - (x_1/x_0)^{-\alpha^*}}{\alpha^{*2}}$,
- ξ_2 is a gaussian vector of mean 0 and of variance $\sigma_2^2 = \frac{(x_1/x_0)^{-\alpha^*}}{\alpha^{*2}}$,
- W is a Brownian motion on \mathbb{R} and $c = \left(\frac{(x_1/x_0)^{-\alpha^*}}{\alpha^{*x_1}} \right)^{-1/2}$,
- E is an exponential variable of parameter α^* .

Moreover, ξ_1, ξ_2, W, E are independent.

3. Estimation

For a fixed real number x , if we call $\Upsilon(x)$ (respectively $\overline{\Upsilon(x)}$) the subset of the data larger than x (respectively smaller) and $N(x)$ (respectively $n - N(x)$) its cardinality, then, the maximum likelihood estimators of the parameters $x_0, x_1, \alpha_n, \beta_n$ are defined by the following equalities:

$$\begin{aligned} \hat{x}_0 &= X_{(1)}, \\ \hat{\alpha} &= \left[\frac{1}{n - N(\hat{x}_1)} \sum_{X_i \in \overline{\Upsilon(\hat{x}_1)}} \log \left(\frac{X_i}{\hat{x}_0} \right) + \frac{N(\hat{x}_1)}{n - N(\hat{x}_1)} \log \left(\frac{\hat{x}_1}{\hat{x}_0} \right) \right]^{-1}, \\ \hat{\beta} &= \left[\frac{1}{N(\hat{x}_1)} \sum_{X_i \in \Upsilon(\hat{x}_1)} \log \left(\frac{X_i}{\hat{x}_1} \right) \right]^{-1}, \\ \hat{x}_1 &= \arg \max_{x \in \{X_1, \dots, X_n\}} \sum_{X_i \in \Upsilon(x)} \left(\log \left(\frac{\hat{\beta}}{\hat{\alpha}} \right) + (\hat{\alpha} - \hat{\beta}) \log \left(\frac{X_i}{x} \right) \right). \end{aligned}$$

THEOREM 2. – *Under the same hypotheses as in Theorem 1, the maximum likelihood estimators of the four parameters are asymptotically independent and have the following asymptotic distributions (under $\mathbb{P}_{x_0, x_1, \alpha_n, \beta_n}^n$):*

$$\begin{cases} n(\hat{x}_0 - x_0) & \rightarrow \text{Exp}(\alpha^*) \\ \sqrt{n}(\hat{\alpha} - \alpha_n) & \rightarrow \mathcal{N}(0, \sigma_1^2) \\ \sqrt{n}(\hat{\beta} - \beta_n) & \rightarrow \mathcal{N}(0, \sigma_2^2) \\ nd_n^2(\hat{x}_1 - x_1) & \rightarrow \zeta \end{cases}$$

where $\zeta = \arg \max_{t \in \mathbb{R}} c(W(t) - \frac{|t|}{2})$ and $\sigma_1^2, \sigma_2^2, c$ are defined in Theorem 1.

Table 1

	n_1	n_2	x_0	x_1	α	β
True			10	20	2	2.1
Empirical mean	1524	1235	10.0021	20.0182	1.8145	2.0929
Empirical s.d.			0.0000	0.0068	0.0005	0.0042

Table 2

	n_1	n_2	x_0	x_1	α	β
True			10	20	2.1	2
Empirical mean	610	213	10.0059	21.0152	2.0000	2.7930
Empirical s.d.			0.0000	0.0079	0.0534	17.8443
True			10	20	2.1	2
Empirical mean	163	704	10.0265	19.9967	0.2938	1.9954
Empirical s.d.			0.0006	0.0030	0.0002	0.0053
True			10	20	2.1	2
Empirical mean	60	1167	10.0509	19.9527	0.0757	1.9972
Empirical s.d.			0.0035	0.0123	0.0000	0.0044
True			10	20	2	2.1
Empirical mean	1520	64	10.0028	10.9994	2.4418	3.5901
Empirical s.d.			0.0000	1.0272	0.0320	0.1043

3.1. Simulation results

For each simulation, we work with $N = 50$ samples of the data; we give the number n_1 (respectively n_2) of data falling between x_0 and x_1 (respectively larger than x_1). For every simulation, we observe that x_0 is the parameter which is the best estimated. It was expected since its rate of convergence is the fastest. In the opposite, the rate d_n is the smallest one and then, the empirical variance of \hat{x}_1 is the worse one.

First, we consider a large number of data. The results are excellent (see Table 1): our procedure is able to detect a very small break between α and β .

Next, we consider the case where n_1 is varying but we keep relatively large values for n_2 . As expected, small values of n_1 , seriously damage the estimation of α , however, the other parameter are encouragingly well fitted (see Table 2). For small n_2 , the procedure does not detect well the change point even if n_1 is large. In fact, we find here a problem quite similar to the problem of the “horror plot” when we compute the Hill estimator of the tail parameter. The likelihood (as function of x) is very erratic for small x and then

Table 3

	n_1	n_2	x_0	x_1	α	β
True			10	15	1.5	2.1
Empirical mean	658	55	10.0047	11.3303	3.0660	4.0971
Empirical s.d.			0.0000	3.1229	0.3824	0.9730
True			10	20	1.5	2.1
Empirical mean	630	68	10.0065	11.1964	1.9784	2.9256
Empirical s.d.			0.0000	0.6113	0.0330	0.0327
True			10	50	1.5	2.1
Empirical mean	641	83	10.0087	35.8182	1.4297	2.0191
Empirical s.d.			0.0000	722.5523	0.0338	0.4354

x_1 is always under-estimated. This problem does not appear when n_2 is large (even if n_1 is small).

Table 3 more precisely investigates the case of small n_2 . We observe that results are better when $x_1 - x_0$ is large enough for the parameters α and β . In the first lines, the true parameters α and β do not belong to the interval $[\text{emp.mean} - 2\text{emp.s.d.}, \text{emp.mean} + 2\text{emp.s.d.}]$. But in the last case, the change point is not well estimated (see the standard error).

4. Proofs

4.1. Proof of Theorem 1

The proof, as usual, relies on two classical arguments. The first one consists in proving that the finite distributions are converging. This is an elementary consequence of Proposition 1 (see Section 4.2).

The second argument is generally somewhat more complex. It consists in proving the tightness of the sequence. As can be seen in Proposition 1, the log-likelihood splits into three terms. The two first ones are only involving standard arguments to prove the tightness. For their behavior we refer to Ibragimov and Khas'minskii [7] or Deshayes and Picard [3]. However, the last term requests a more careful attention; its behaviour is studied in Proposition 2 (see Section 4.3).

In the sequel, we omit the parameters in the expectations and the probabilities, but they are taken without ambiguity under the distribution $\mathbb{P}_{x_0, x_1, \alpha_n, \beta_n}^n$.

4.2. Expansion of the likelihood ratio

PROPOSITION 1. – *Let t, u, v, w be some constants and assume that $\lim_{n \rightarrow +\infty} nd_n^4 = +\infty$ and $\lim_{n \rightarrow +\infty} d_n = 0$. Then, we have*

$$\log \Gamma_n(u, v, w, t) = \left[\left(-u\xi_{1,n} - \frac{1}{2}u^2\sigma_1^2 \right) + \left(-v\xi_{2,n} - \frac{1}{2}v^2\sigma_2^2 \right) + \left(d_n W_n(t) - \frac{\sigma^2(t)}{2} \right) + w \frac{\alpha^*}{x_0} + o_P(1) \right] \mathbf{1}_{\{E_n > w + o(1)\}}$$

where

$$\begin{aligned} \xi_{1,n} &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \left[\left(Y_i - \frac{1}{\alpha_n} \right) \mathbf{1}_{\{\log(\frac{x_0}{x_1} + \frac{w}{nx_1}) \leq Y_i \leq 0\}} - \mathbb{E} \left(Y_i - \frac{1}{\alpha_n} \right) \mathbf{1}_{\{\log(\frac{x_0}{x_1} + \frac{w}{nx_1}) \leq Y_i \leq 0\}} \right] \mathbf{1}_{\{t > 0\}} \\ &\quad + \frac{1}{\sqrt{n}} \sum_{i=1}^n \left[\left(Y_i - \frac{1}{\alpha_n} \right) \mathbf{1}_{\{\log(\frac{x_0}{x_1} + \frac{w}{nx_1}) \leq Y_i \leq \log(1 + \frac{t}{x_1 n d_n^2})\}} \right. \\ &\quad \left. - \mathbb{E} \left(Y_i - \frac{1}{\alpha_n} \right) \mathbf{1}_{\{\log(\frac{x_0}{x_1} + \frac{w}{nx_1}) \leq Y_i \leq \log(1 + \frac{t}{x_1 n d_n^2})\}} \right] \mathbf{1}_{\{t < 0\}}, \\ \xi_{2,n} &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \left[\left(Z_i - \frac{1}{\beta_n} \right) \mathbf{1}_{\{\log(1 + \frac{t}{x_1 n d_n^2}) \leq Z_i\}} - \mathbb{E} \left(Z_i - \frac{1}{\beta_n} \right) \mathbf{1}_{\{\log(1 + \frac{t}{x_1 n d_n^2}) \leq Z_i\}} \right] \mathbf{1}_{\{t > 0\}} \\ &\quad + \frac{1}{\sqrt{n}} \sum_{i=1}^n \left[\left(Z_i - \frac{1}{\beta_n} \right) \mathbf{1}_{\{0 \leq Z_i\}} - \mathbb{E} \left(Z_i - \frac{1}{\beta_n} \right) \mathbf{1}_{\{0 \leq Z_i\}} \right] \mathbf{1}_{\{t < 0\}}, \\ W_n(t) &= \sum_{i=1}^n \left[\left(Z_i - \frac{1}{\beta_n} \right) \mathbf{1}_{\{0 \leq Z_i \leq \log(1 + \frac{t}{x_1 n d_n^2})\}} - \mathbb{E} \left(Z_i - \frac{1}{\beta_n} \right) \mathbf{1}_{\{0 \leq Z_i \leq \log(1 + \frac{t}{x_1 n d_n^2})\}} \right] \mathbf{1}_{\{t > 0\}} \\ &\quad + \sum_{i=1}^n \left[\left(Z_i - \frac{1}{\beta_n} \right) \mathbf{1}_{\{\log(1 + \frac{t}{x_1 n d_n^2}) \leq Z_i \leq 0\}} - \mathbb{E} \left(Z_i - \frac{1}{\beta_n} \right) \mathbf{1}_{\{\log(1 + \frac{t}{x_1 n d_n^2}) \leq Z_i \leq 0\}} \right] \mathbf{1}_{\{t < 0\}}, \\ E_n &= n \log \frac{X(1)}{x_0}. \end{aligned}$$

The variables Y_i, Z_i are

$$Y_i = \log \frac{X_i}{x_1} \mathbf{1}_{\{x_0 \leq X \leq x_1\}}, \quad Z_i = \log \frac{X_i}{x_1} \mathbf{1}_{\{x_1 \leq X\}} \tag{2}$$

and

$$\begin{aligned} \sigma_1^2 &= \frac{1 - (x_1/x_0)^{-\alpha^*}}{\alpha^{*2}} = V(\xi_{1,n}) + o(1), \\ \sigma_2^2 &= \frac{(x_1/x_0)^{-\alpha^*}}{\alpha^{*2}} = V(\xi_{2,n}) + o(1), \\ \sigma^2(t) &= \frac{(x_1/x_0)^{-\alpha^*}}{\alpha^* x_1} |t| = V(d_n W_n(t)) + o(1). \end{aligned}$$

Proof of Proposition 1. – We only give the proof for $t > 0$. The other case is similar, for reason of symmetry. Let $h_{0,n}, h_{1,n}, \eta_n$ and δ_n be sequences tending to zero: in the sequel, we omit the subscript n .

The likelihood of the sample is

$$L_n(x_0, x_1, \alpha, \beta, X_1, \dots, X_n) = 1_{\{x_0 \leq X_{(1)}\}} \times \prod_{i=1}^n (\alpha x_0^\alpha X_i^{-(\alpha+1)})^{1_{\{x_0 \leq X_i \leq x_1\}}} \\ \times \prod_{i=1}^n (\beta x_1^{\beta-\alpha} x_0^\alpha X_i^{-(\beta+1)})^{1_{\{x_1 \leq X_i\}}}$$

where $X_{(1)} = \min(X_1, \dots, X_n)$. The Log-likelihood ratio splits into three terms:

$$\log \frac{L_n(x_0 + h_0, x_1 + h_1, \alpha + \eta, \beta + \delta)}{L_n(x_0, x_1, \alpha, \beta)} = (T_1 + T_2 + T_3) (1 + \log 1_{\{x_0+h_0 < X_{(1)}\}}). \quad (3)$$

We have

$$T_1 = \sum_{i=1}^n 1_{\{x_0+h_0 \leq X_i \leq x_1\}} \left[\log \left(1 + \frac{\eta}{\alpha} \right) + (\alpha + \eta) \log \left(1 + \frac{h_0}{x_0} \right) + \eta \log \frac{x_0}{x_1} - \eta \log \frac{X_i}{x_1} \right],$$

$$T_2 = \sum_{i=1}^n 1_{\{x_1 \leq X_i \leq x_1+h_1\}} \left[\log \left(1 - \frac{\beta - \alpha - \eta}{\beta} \right) + (\alpha + \eta) \log \left(1 + \frac{h_0}{x_0} \right) \right. \\ \left. + \eta \log \frac{x_0}{x_1} + (\beta - \alpha - \eta) \log \frac{X_i}{x_1} \right],$$

$$T_3 = \sum_{i=1}^n 1_{\{x_1+h_1 \leq X_i\}} \left[\log \left(1 + \frac{\delta}{\beta} \right) + (\beta - \alpha + \delta - \eta) \log \left(1 + \frac{h_1}{x_1} \right) \right. \\ \left. + (\alpha + \eta) \log \left(1 + \frac{h_0}{x_0} \right) - \eta \log \frac{x_1}{x_0} - \delta \log \frac{X_i}{x_1} \right].$$

Let us set

$$I = 1_{\{\log(\frac{x_0+h_0}{x_1}) \leq Y \leq 0\}}, \quad J = 1_{\{0 \leq Z \leq \log(1+\frac{h_1}{x_1})\}}, \quad K = 1_{\{\log(1+\frac{h_1}{x_1}) \leq Z\}}.$$

and recall the change for variables (2):

$$Y = \log(X/x_1) 1_{\{x_0 \leq X \leq x_1\}} \quad \text{and} \quad Z = \log(X/x_1) 1_{\{x_1 \leq X\}}.$$

We obtain

$$T_1 = -\eta \sum_{i=1}^n \left[\left(Y_i - \frac{1}{\alpha} \right) I_i - \mathbb{E} \left(Y_i - \frac{1}{\alpha} \right) I_i \right] - \eta n \mathbb{E} \left(Y - \frac{1}{\alpha} \right) I \\ + \left[-\frac{\eta}{\alpha} + \log \left(1 + \frac{\eta}{\alpha} \right) + (\alpha + \eta) \log \left(1 + \frac{h_0}{x_0} \right) - \eta \log \left(\frac{x_0}{x_1} \right) \right] \sum_{i=1}^n (I_i - \mathbb{E} I_i) \\ + n \left[-\frac{\eta}{\alpha} + \log \left(1 + \frac{\eta}{\alpha} \right) + (\alpha + \eta) \log \left(1 + \frac{h_0}{x_0} \right) - \eta \log \left(\frac{x_0}{x_1} \right) \right] \mathbb{E} I,$$

$$T_2 = -(\alpha - \beta + \eta) \sum_{i=1}^n \left[\left(Z_i - \frac{1}{\beta} \right) J_i - \mathbb{E} \left(Z_i - \frac{1}{\beta} \right) J_i \right] - (\alpha - \beta + \eta) n \mathbb{E} \left(Z - \frac{1}{\beta} \right) J \\ + \left[-\frac{\alpha - \beta + \eta}{\beta} + \log \left(1 - \frac{\beta - \alpha + \eta}{\beta} \right) + (\alpha + \eta) \log \left(1 + \frac{h_0}{x_0} \right) \right]$$

$$\begin{aligned}
 & + \eta \log\left(\frac{x_0}{x_1}\right) \sum_{i=1}^n (J_i - \mathbb{E}J_i) + n \left[-\frac{\alpha - \beta + \eta}{\beta} + \log\left(1 - \frac{\beta - \alpha + \eta}{\beta}\right) \right. \\
 & \left. + (\alpha + \eta) \log\left(1 + \frac{h_0}{x_0}\right) + \eta \log\left(\frac{x_0}{x_1}\right) \right] \mathbb{E}J, \\
 T_3 = & -\delta \sum_{i=1}^n \left[\left(Z_i - \frac{1}{\beta} \right) K_i - \mathbb{E}\left(Z_i - \frac{1}{\beta} \right) K_i \right] - \delta n \mathbb{E}\left(Z - \frac{1}{\beta} \right) K \\
 & + \left[-\frac{\delta}{\beta} + \log\left(1 + \frac{\delta}{\beta}\right) + (\beta + \delta) \log\left(1 + \frac{h_0}{x_0}\right) - \eta \log\left(\frac{x_0}{x_1}\right) \right] \sum_{i=1}^n (K_i - \mathbb{E}K_i) \\
 & + n \left[-\frac{\delta}{\beta} + \log\left(1 + \frac{\delta}{\beta}\right) + (\beta + \delta) \log\left(1 + \frac{h_0}{x_0}\right) - \eta \log\left(\frac{x_0}{x_1}\right) \right] \mathbb{E}K.
 \end{aligned}$$

Using the distribution function (1) of X , we easily compute the expectation of the variables of interest

$$\begin{aligned}
 \mathbb{E}I &= \left(\frac{x_1}{x_0}\right)^{-\alpha} \left[\left(\frac{x_0}{x_1} + \frac{h_0}{x_1}\right)^{-\alpha} - 1 \right], \\
 \mathbb{E}J &= \left(\frac{x_1}{x_0}\right)^{-\alpha} \left[1 - \left(1 + \frac{h_1}{x_1}\right)^{-\beta} \right], \\
 \mathbb{E}K &= \left(\frac{x_1}{x_0}\right)^{-\alpha} \left(1 + \frac{h_1}{x_1}\right)^{-\beta}, \\
 \mathbb{E}\left(Y - \frac{1}{\alpha}\right)I &= \log\left(\frac{x_0}{x_1} + \frac{h_0}{x_1}\right) \left(1 + \frac{h_0}{x_0}\right)^{-\alpha}, \\
 \mathbb{E}\left(Z - \frac{1}{\beta}\right)J &= \left(\frac{x_1}{x_0}\right)^{-\alpha} \left[-\log\left(1 + \frac{h_1}{x_1}\right) \left(1 + \frac{h_1}{x_1}\right)^{-\beta} \right], \\
 \mathbb{E}\left(Z - \frac{1}{\beta}\right)K &= \left(\frac{x_1}{x_0}\right)^{-\alpha} \left(1 + \frac{h_1}{x_1}\right)^{-\beta} \log\left(1 + \frac{h_1}{x_1}\right).
 \end{aligned}$$

Replacing now

$$\eta = \frac{u}{\sqrt{n}}, \quad \delta = \frac{v}{\sqrt{n}}, \quad h_0 = \frac{w}{n}, \quad h_1 = \frac{t}{nd_n^2}$$

and assuming the conditions $\lim_{n \rightarrow +\infty} nd_n^4 = +\infty$ and $\lim_{n \rightarrow +\infty} d_n = 0$, it follows

$$\begin{aligned}
 T_1 &= -u\xi_{1,n} + u\sqrt{n} \log\left(\frac{x_0}{x_1}\right) \left(\frac{x_0}{x_1}\right)^{\alpha^*} - u^2 \frac{1 - \left(\frac{x_0}{x_1}\right)^{\alpha^*}}{2\alpha^{*2}} \\
 & \quad + \frac{\alpha^* w}{x_0} \left(1 - \left(\frac{x_0}{x_1}\right)^{\alpha^*}\right) + o(1) + o_P(1), \\
 T_2 &= -d_n W_n(t) - \frac{t}{2} \left(\frac{x_0}{x_1}\right)^{\alpha^*} \frac{1}{\beta x_1} - \frac{t}{d_n} \left(\frac{x_0}{x_1}\right)^{\alpha^*} \frac{1}{x_1} + o(1) + o_P(1),
 \end{aligned}$$

$$T_3 = -v\xi_2 - u\sqrt{n} \log\left(\frac{x_0}{x_1}\right) \left(\frac{x_0}{x_1}\right)^{\alpha^*} - v^2 \frac{1}{2\beta^2} \left(\frac{x_0}{x_1}\right)^{\alpha^*} + \frac{\alpha^* w}{x_0} \left(\frac{x_0}{x_1}\right)^{\alpha^*} + \frac{t}{d_n} \left(\frac{x_0}{x_1}\right)^{\alpha^*} \frac{1}{x_1} + o(1) + o_P(1).$$

Since $1_{\{x_0+h_0 \leq X_{(1)}\}} = 1_{\{w+o(1) \leq n \log \frac{X_{(1)}}{x_0}\}}$, combining with (3), we obtain the result.

4.3. Convergence of the process $\{d_n W_n(t), t > 0\}$

Let us recall that the process $W_n(t)$ is defined, for $t > 0$, by

$$W_n(t) = \sum_{i=1}^n \left[\left(Z_i - \frac{1}{\beta_n} \right) 1_{\{0 \leq Z_i \leq \log(1 + \frac{t}{x_1 n d_n^2})\}} - E \left(Z_i - \frac{1}{\beta_n} \right) 1_{\{0 \leq Z_i \leq \log(1 + \frac{t}{x_1 n d_n^2})\}} \right]$$

where $Z = \log(X/x_1) 1_{\{x_1 \leq X\}}$ and that $V(d_n W_n(t)) = t \frac{(x_0/x_1)^{\alpha^*}}{\alpha^*} (1 + o(1))$.

PROPOSITION 2. – *We assume that $\lim_{n \rightarrow +\infty} \alpha_n = \alpha^*$, $\lim_{n \rightarrow +\infty} n d_n^4 = +\infty$ and $\lim_{n \rightarrow +\infty} d_n = 0$. Let us denote $\{d_n \tilde{W}_n(t), t > 0\}$ the linear interpolation of $\{d_n W_n(t), t > 0\}$. The process $\{d_n \tilde{W}_n(t), t > 0\}$ weakly converges under \mathbb{P} to the process $\{cW(t), t > 0\}$ with the topology of uniform convergence on compact sets for $c = \sqrt{\frac{\alpha^* x_1}{(x_0/x_1)^{\alpha^*}}}$.*

Proof of Proposition 2. – The first argument of the proof consists in proving the convergence of the finite distributions and is an elementary consequence of the central limit theorem.

We concentrate on proving the tightness of the sequence. Using standard arguments, it is enough to prove that for $0 < \tau < 1/2$, for any $T > 0$, $\varepsilon > 0$, there exists $\Lambda > 0$ such that

$$\forall n \geq 1, \quad \mathbb{P}(\|d_n \tilde{W}_n\|_{\tau, T} \geq \Lambda) \leq \varepsilon, \tag{4}$$

where $\|\cdot\|_{\tau, T}$ denotes the Lipschitz-norm,

$$\|f\|_{\tau, T} = |f(0)| + \sup_{x \neq y \in [0, T]} \frac{|f(x) - f(y)|}{|x - y|^\tau}. \tag{5}$$

To prove (4), we use an equivalent form of the Lipschitz-norm using the Ciesielski basis. If $\{\chi_{j,k}, j \geq 1, k \in \mathbb{N}\}$ is the Haar basis, we define the Ciesielski basis as follows:

$$\begin{aligned} \Delta_{-10}(t) &= 1_{[0, T]}(t), & \Delta_{00}(t) &= \frac{t}{T} 1_{[0, T]}(t), \\ \Delta_{jk}(t) &= \int_0^{\frac{t}{T}} \chi_{j,k}(u) du, & j \geq 1, k \in \mathbb{N}. \end{aligned}$$

This basis has the very useful following properties (see Ciesielski [1], and notice also the use of such a property in a stochastic framework in Kerkycharian, Roynette [6]):

1. If $f = \sum_{j,k} \lambda_{j,k} \Delta_{j,k}$ is a continuous function on $[0, T]$, its coefficients satisfy the relations:

$$\begin{aligned} \lambda_{-1,0} &= f(0), & \lambda_{0,0} &= f(T) - f(0), \\ \lambda_{j,k} &= f\left(\frac{(k+1/2)T}{2^j}\right) - \frac{1}{2} \left[f\left(\frac{kT}{2^j}\right) + f\left(\frac{(k+1)T}{2^j}\right) \right]. \end{aligned}$$

2. The lip-norm introduced in (5) is equivalent to the following norm

$$\|f\|'_{\tau,T} = \sup_{j \geq -1,k} 2^{j\tau} |\lambda_{j,k}|. \tag{6}$$

We use the points 1 and 2 to prove (4). More precisely, if we consider:

$$\begin{aligned} \tilde{\lambda}_{-1,0}^n &= d_n \tilde{W}_n(0) = 0, \\ \tilde{\lambda}_{0,0}^n &= d_n (\tilde{W}_n(T) - \tilde{W}_n(0)), \\ \tilde{\lambda}_{j,k}^n &= d_n \tilde{W}_n\left(\frac{(k+1/2)T}{2^j}\right) - \frac{d_n}{2} \left(\tilde{W}_n\left(\frac{kT}{2^j}\right) + \tilde{W}_n\left(\frac{(k+1)T}{2^j}\right) \right), \end{aligned}$$

it is enough to prove that:

$$\forall \varepsilon > 0, \exists \Lambda_\varepsilon, \quad \mathbb{P}(\sup_{j,k} 2^{j\tau} |\tilde{\lambda}_{j,k}^n| > \Lambda_\varepsilon) \leq \varepsilon. \tag{7}$$

We have a slightly different strategy concerning the ‘small’ and the ‘large’ j ’s. Let us now fix $\Lambda > 0$ and J_0 such that

$$2^{J_0} = \left(\frac{\Lambda}{6Cd_n}\right)^{1/\tau} \tag{8}$$

for $C = \log(\gamma_1 T + \gamma_2)$ where $\gamma_1 \geq (x_1 n d_n^2)^{-1}$ and $\gamma_2 \geq (\beta_n)^{-1}$. We obviously have:

$$\mathbb{P}(\sup_{j,k} 2^{j\tau} |\tilde{\lambda}_{j,k}^n| > \Lambda) \leq \sum_{j \leq J_0} \mathbb{P}(\sup_k 2^{j\tau} |\tilde{\lambda}_{j,k}^n| > \Lambda) + \sum_{j > J_0} \mathbb{P}(\sup_k 2^{j\tau} |\tilde{\lambda}_{j,k}^n| > \Lambda). \tag{9}$$

4.3.1. First case: $j \leq J_0$

We first remark that the terms corresponding to $j \in \{0, 1\}$ does not present difficulties. For the other ones, we write:

$$\tilde{\lambda}_{j,k}^n = \lambda_{j,k}^n + \varepsilon_{j,k} \tag{10}$$

where $\lambda_{j,k}^n$ is calculated as $\tilde{\lambda}_{j,k}^n$ replacing \tilde{W}_n by W_n (i.e. forgetting the interpolation), and

$$\begin{aligned} |\varepsilon_{j,k}| &\leq 3 \sup_{[0, T/x_1 n d_n^2]} |d_n \tilde{W}_n(t) - d_n W_n(t)| \\ &\leq 3d_n \left| \log\left(\frac{T}{x_1 n d_n^2} - \frac{1}{\beta_n}\right) \right| \\ &\leq 3Cd_n \end{aligned} \tag{11}$$

for C used in the definition (8). Hence, we have:

$$\sum_{j \leq J_0} \mathbb{P}(\sup_k 2^{j\tau} |\tilde{\lambda}_{j,k}^n| > \Lambda) \leq \sum_{j \leq J_0} [\mathbb{P}(\sup_k 2^{j\tau} |\lambda_{j,k}^n| > \Lambda/2) + \mathbb{P}(\sup_k 2^{j\tau} |\varepsilon_{j,k}| > \Lambda/2)]. \tag{12}$$

Using (8) and (11), we observe that the second part of the sum is null. For the first part, using the definition of the coefficients $\lambda_{j,k}^n$, we have:

$$\begin{aligned} \forall j \geq 1, \quad & \mathbb{P}(\sup_k 2^{j\tau} |\lambda_{j,k}^n| > \Lambda/2) \\ & \leq 2\mathbb{P}\left(\sup_k 2^{j\tau} d_n \left| W_n\left(\frac{(k+1)T}{2^j}\right) - W_n\left(\frac{kT}{2^j}\right) \right| > \Lambda/2\right) \\ & \leq 2\mathbb{P}(\sup_k 2^{j\tau} d_n |\Delta W_n(j, k)| > \Lambda/2) := 2R \end{aligned} \tag{13}$$

where, for arbitrary j, k ,

$$\begin{aligned} \Delta W_n(j, k) &= W_n\left(\frac{(k+1)T}{2^j}\right) - W_n\left(\frac{kT}{2^j}\right) \\ &= \sum_{i=1}^n \left[\left(Z_i - \frac{1}{\beta_n}\right) L_i - \mathbb{E}\left(Z_i - \frac{1}{\beta_n}\right) L_i \right] \end{aligned}$$

for $L_i = 1_{\{\log(1 + \frac{kT}{x_1 2^j n d_n^2}) \leq Z_i \leq \log(1 + \frac{(k+1)T}{x_1 2^j n d_n^2})\}}$. Remembering the distribution (1) of the X_i 's, elementary computation gives the variance of $\Delta W_n(j, k)$:

$$V[\Delta W_n(j, k)] = \frac{1}{\beta_n} \left(\frac{x_0}{x_1}\right)^\alpha \frac{T}{x_1 2^j d_n^2} (1 + o(1)). \tag{14}$$

We shall now use a standard argument in the context of empirical processes (see for instance Pollard [9]). Let us consider \mathfrak{S}_k the sigma field generated by the variables $\{\Delta W_n(j, u), u \leq k\}$ for n and $j \geq 1$ fixed. We are going to take advantage of Lemma 1 (see below). Its proof is rejected at the end of this subsection.

LEMMA 1. – For Λ large enough, uniformly for $j \geq 0$, on the event

$$|\Delta W_n(j, k)| \geq \frac{\Lambda}{2^{j\tau+1} d_n},$$

we have

$$\mathbb{P}(|\Delta W_n(j, (2^j - 1)T) - \Delta W_n(j, k)| \leq |\Delta W_n(j, k)|/2 | \mathfrak{S}_k) \geq 1/2. \tag{15}$$

We apply Lemma 1 by introducing the stopping time:

$$s = \inf\{k \in \mathbb{N}, |\Delta W_n(j, k)| > \Lambda 2^{-(j\tau+1)} d_n^{-1}\}.$$

As usual, we put $s = \infty$ if the event never occurs. Since the event $\{s = k\}$ is \mathfrak{S}_k -measurable, the right side term R of (13) is bounded as follows:

$$\begin{aligned}
 R &\leq \sum_{0 \leq k < T2^j} \mathbb{P}(s = k) = \sum_{0 \leq k < T2^j} \mathbb{E}1_{\{s=k\}} \\
 &\leq 2 \sum_{0 \leq k < T2^j} \mathbb{E}1_{\{s=k\}} \mathbb{P}(|\Delta W_n(j, (2^j - 1)T) - \Delta W_n(j, k)| \leq |\Delta W_n(j, k)/2| / \mathfrak{S}_k) \\
 &= 2 \sum_{0 \leq k < T2^j} \mathbb{E}1_{\{s=k\}} \mathbb{E}[1_{\{|\Delta W_n(j, (2^j - 1)T) - \Delta W_n(j, k)| \leq |\Delta W_n(j, k)/2|\}} / \mathfrak{S}_k] \\
 &= 2 \sum_{0 \leq k < T2^j} \mathbb{E}1_{\{s=k\}} 1_{\{|\Delta W_n(j, (2^j - 1)T) - \Delta W_n(j, k)| \leq |\Delta W_n(j, k)/2|\}}.
 \end{aligned}$$

Hence, using Chebychev inequality and the bound (14) on the variance of ΔW_n , we get:

$$\begin{aligned}
 R &\leq \sum_{0 \leq k < T2^j} \mathbb{P}(|\Delta W_n(j, (2^j - 1)T) - \Delta W_n(j, k)| \leq |\Delta W_n(j, k)/2| \cap s = k) \\
 &\leq \sum_{0 \leq k < T2^j} \mathbb{P}(|\Delta W_n(j, (2^j - 1)T)| \geq \Lambda 2^{-(j\tau+2)} d_n^{-1} \cap s = k) \\
 &\leq \mathbb{P}(|\Delta W_n(j, (2^j - 1)T)| \geq \Lambda 2^{-(j\tau+2)} d_n^{-1}) \\
 &\leq c \Lambda^{-2} 2^{j(2\tau-1)}
 \end{aligned} \tag{16}$$

for $c = 2^5 (\frac{x_0}{x_1})^{\alpha^*} \frac{1}{\alpha^*} \frac{T}{x_1}$. Combining (12), (13), (16), we get:

$$\sum_{0 \leq j \leq J_0} \mathbb{P}(\sup_k 2^{j\tau} |\tilde{\lambda}_{j,k}^n| > \Lambda) \leq c \Lambda^{-2} \sum_{j \geq 0} 2^{j(2\tau-1)}.$$

Since $\tau < 1/2$ and choosing Λ large enough, this quantity may be bounded (independently of n) by an arbitrary small quantity.

4.3.2. Second case $j > J_0$

For this case, we follow the same steps as for the previous one, but some modifications are necessary:

1. First we replace as in (10), \tilde{W}_n by W_n . However, the error may be bounded in the following way: we take benefit of the following remark. If $W_n(\frac{kT}{2^j}) = W_n(\frac{(k+1)T}{2^j})$, then $\varepsilon_{j,k} = 0$ since the $\lambda_{j,k}$'s for $j > 0$ do not 'see' the straight lines. Moreover, we have:

$$|\varepsilon_{j,k}| \leq r \left| d_n U_n \left(\frac{kT}{2^j} \right) - d_n U_n \left(\frac{(k+1)T}{2^j} \right) \right| \tag{17}$$

for $r = \gamma \log(1 + T)$ where $\gamma > 0$ is a constant and

$$U_n(t) = \sum_{i=1}^n 1_{\{0 \leq Z_i \leq \log(1+t/x_1 n d_n^2)\}}. \tag{18}$$

2. Hence we deduce, using (16) and (17):

$$\begin{aligned}
 &\sum_{j > J_0} \mathbb{P}(\sup_k 2^{j\tau} |\tilde{\lambda}_{j,k}^n| > \Lambda) \\
 &\leq \sum_{j > J_0} \mathbb{P}\left(\sup_k 2^{j\tau} r \left| d_n U_n \left(\frac{kT}{2^j} \right) - d_n U_n \left(\frac{(k+1)T}{2^j} \right) \right| > \Lambda/2\right)
 \end{aligned}$$

$$\begin{aligned}
 &+ \sum_{j \geq 0} \mathbb{P}(\sup_k 2^{j\tau} |\lambda_{j,k}^n| > \Lambda/2) \\
 &\leq \sum_{j > J_0} \mathbb{P}(\sup_k 2^{j\tau} r d_n |\Delta U_n(j, k)| > \Lambda/2) + c2^5 \Lambda^{-2} \sum_{j \geq 0} 2^{j(2\tau-1)}, \tag{19}
 \end{aligned}$$

where

$$\Delta U_n(j, k) = U_n\left(\frac{(k+1)T}{2^j}\right) - U_n\left(\frac{kT}{2^j}\right) = \sum_{i=1}^n L_i$$

for L_i defined in the previous subsection.

3. The next step consists in bounding the first sum. We use the same arguments as above replacing W_n by U_n . Using now the definition (8) of J_0 and thanks to the hypotheses on d_n , we get

$$\begin{aligned}
 \mathbb{E}2^{j\tau} |\Delta U_n(j, k)| &= (x_1/x_0)^{\alpha^*} \frac{\alpha^* T}{x_1} \frac{2^{j(\tau-1)}}{nd_n^2} (1 + o(1)) \\
 &\leq \Lambda d_n^{-1} / 4
 \end{aligned}$$

choosing $\Lambda \geq [4(x_1/x_0)^{\alpha^*} (\alpha^* T/x_1)]^\tau (6C)^{1-\tau}$. This remark allows us to consider

$$\Delta U_n^*(j, k) = \Delta U_n(j, k) - \mathbb{E}[\Delta U_n(j, k)]$$

and to get

$$\mathbb{P}\left(\sup_{0 \leq k \leq T2^j} 2^{j\tau} r d_n |\Delta U_n(j, k)| > \Lambda/2\right) \leq \mathbb{P}\left(\sup_{0 \leq k \leq rT2^j} 2^{j\tau} r d_n |\Delta U_n^*(j, k)| > \Lambda/4\right).$$

4. We can prove a complete analogue of Lemma 1, replacing ΔW_n by ΔU_n^* .
5. Using exactly the same arguments as for (16), we deduce:

$$\begin{aligned}
 \mathbb{P}\left(\sup_{0 \leq k \leq rT2^j} 2^{j\tau} r |\Delta U_n^*(j, k)| > \Lambda/4\right) &\leq \mathbb{P}(r |\Delta U_n^*(j, (2^j - 1)T)| \geq \Lambda 2^{-(j\tau+2)}/2) \\
 &\leq c' r^2 \Lambda^{-2} 2^{-j(1-2\tau)}.
 \end{aligned}$$

4.3.3. Proof of Lemma 1

$$\begin{aligned}
 &\Delta W_n(j, (2^j - 1)T) - \Delta W_n(j, k) \\
 &= \left(W_n(T) - W_n\left(\frac{(2^j - 1)T}{2^j}\right)\right) - \left(W_n\left(\frac{(k+1)T}{2^j}\right) - W_n\left(\frac{kT}{2^j}\right)\right) \\
 &= \sum_{i=1}^n \left(Z_i - \frac{1}{\beta}\right) 1_{\{\log(1 + \frac{T}{x_1 n d_n^2} (1 - \frac{1}{2^j})) \leq Z_i \leq \log(1 + \frac{T}{x_1 n d_n^2})\}} \\
 &\quad - \sum_{i=1}^n \left(Z_i - \frac{1}{\beta}\right) 1_{\{\log(1 + \frac{Tk}{x_1 n d_n^2}) \leq Z_i \leq \log(1 + \frac{T(k+1)}{x_1 n d_n^2})\}} - e_n(j, k, T),
 \end{aligned}$$

where

$$e_n(j, k, T) = n\mathbb{E}\left(Z - \frac{1}{\beta}\right) \left[\mathbb{1}_{\{\log(1 + \frac{T}{x_1 n d_n^2} (1 - \frac{1}{2^j})) \leq Z \leq \log(1 + \frac{T}{x_1 n d_n^2})\}} - \mathbb{1}_{\{\log(1 + \frac{Tk}{x_1 n d_n^2}) \leq Z \leq \log(1 + \frac{T(k+1)}{x_1 n d_n^2})\}} \right].$$

Let us remark that the knowledge of $\Delta W_n(j, k)$ implies the knowledge of the number n_z of variables Z_i which belong to $[0, \frac{Tk}{x_1 2^j n d_n^2}]$. Hence, conditionally to \mathfrak{S}_k ,

$$\Delta W_n(j, (2^j - 1)T) - \Delta W_n(j, k) = \sum_{i=n_z+1}^n \tilde{Z}_i$$

where the \tilde{Z}_i 's are i.i.d. centered random variables with variance bounded with $c2^{-j}d_n^{-2}$ where c is depending on $T, x_0, x_1, \alpha, \beta$ (see (14)). Using Bienaymé–Chebychev inequality, the fact that we restrict ourselves to the event $|\Delta W_n(j, k)| \geq \frac{\Lambda}{2^{j\tau+1}d_n}$ and the hypothesis on the sequence d_n , we obtain

$$\mathbb{P}(|\Delta W_n(j, (2^j - 1)T) - \Delta W_n(j, k)| \geq |\Delta W_n(j, k)|/2 | \mathfrak{S}_k) \leq \frac{(n - n_z)V(\tilde{Z}_i)}{(\Delta W_n(j, k)/2)^2} \leq c'2^{j(2\tau-1)}\Lambda^{-2}.$$

Choosing $\Lambda \geq (2c')^{1/2}$ ends the proof.

4.4. Proof of Theorem 2

We only give a sketch of proof of this part. It is a consequence of Theorem 1, and follows rather directly the arguments of Deshayes and Picard [4]. It consists in observing that the Log likelihood process splits into four parts, each corresponding to one of the parameter x_0, x_1, α, β . The part corresponding to x_0 is clear; the parts corresponding to α and β are standard and their behaviour can be handled as in Ibragimov and Khas'minskii [7], for a regular parameter. The only unusual part concerns the term which leads to the argument of the supremum of a shifted Brownian motion. Since we proved the weak convergence with the topology of uniform convergence on compact sets, we have to prove that the arg max in t of $d_n W_n(t) - \sigma^2(t)/2$ necessarily concentrates on compact sets. A way to do that is to prove that the supremum of the process outside $[0, T_0]$ (symmetrically $[-T_0, 0]$), tends to $-\infty$. This is the purpose of the following lines.

As $\sigma^2(t)$ is proportional to $|t|$, a way to attain our goal will be to prove that

$$\mathbb{P}\left(\sup_{|t| \in [T_0, \infty[} \left| \frac{d_n W_n(t)}{t} \right| \geq \varepsilon\right)$$

is arbitrarily small. Of course for reason of symmetry, we will only consider $\mathbb{P}(\sup_{t \in [T_0, \infty[} |d_n W_n(t)/t| \geq \varepsilon)$.

A careful look at the previous proof shows that the constants are coming from the bounding of $|Z_i|$, so in fact we proved that:

$$\mathbb{P}(\|d_n W_n\|_{\tau, [0, T]} \geq \Lambda) \leq AT(1 + \log_2 T)\Lambda^{-2} \quad (20)$$

where now A is an absolute constant, for $\Lambda \geq BT^\tau$, where B is another absolute constant. It follows from (20), using the definition of the Lipschitz norm and the fact that $W_n(0) = 0$:

$$\begin{aligned} \mathbb{P}\left(\sup_{t \in [T_0, \infty[} \left| \frac{d_n W_n(t)}{t} \right| t \geq \varepsilon\right) &\leq \sum_{\log_2(T_0) \leq J} \mathbb{P}\left(\sup_{t \in [2^{J-1}, 2^J]} \left| \frac{d_n W_n(t)}{2^J} \right| \geq \varepsilon\right) \\ &\leq \sum_{\log_2(T_0) \leq J} \mathbb{P}(2^{J\tau} \|d_n W_n\|_{\tau, [0, 2^J]} \geq \varepsilon 2^J) \\ &\leq \sum_{\log_2(T_0) \leq J} A(J+1)2^{-J(1-2\tau)} \varepsilon^{-2} \\ &\leq A' \varepsilon^{-2} \log_2(T_0) T_0^{-(1-2\tau)}. \end{aligned} \quad (21)$$

Note here that the inequality $\tau < 1/2$ was essential, for the finiteness of the serie, but also to ensure that $\Lambda = \varepsilon 2^{J(1-\tau)} \geq B 2^{J\tau}$ and so we can apply (20). Now that we have proved that the argument of the maximum of the process necessarily concentrates on compact sets, it is standard to obtain the limiting result.

REFERENCES

- [1] Z. Ciesielski, On the isomorphisms of the spaces H_α and m , Bull. Acad. Polon. Sci. Ser. Sci. Math. Astronom. Phys. 8 (1960) 217–222.
- [2] L. De Haan, L. Peng, Comparison of tail index estimators, Statistica Neerlandica 52 (1) (1998) 60–70.
- [3] J. Deshayes, D. Picard, Principe d'invariance sur le processus de vraisemblance, Annales de l'I.H.P. 20 (1984) 309–327.
- [4] J. Deshayes, D. Picard, Lois asymptotiques des tests et estimateurs de rupture, Annales de l'I.H.P. 20 (1984) 1–20.
- [5] P. Embrecht, C. Klüppelberg, T. Mikosch, Modelling Extremal Events, Springer-Verlag, Berlin, 1997.
- [6] G. Kerkycharian, B. Roynette, Une démonstration simple des théorèmes de Kolmogorov, Donsker et Ito–Nisio, C. R. Acad. Sci. Paris Ser. I 312 (1991) 877–882.
- [7] I.A. Ibragimov, R.Z. Khas'minskii, Statistical Estimation: Asymptotic Theory, Springer-Verlag, New York, 1981.
- [8] L. LeCam, G. Yang, Asymptotic in Statistics: Some Basic Concepts, Springer-Verlag, New York, 1990.
- [9] D. Pollard, Convergence of Stochastic Processes, Springer-Verlag, New York, 1984.