ASYMPTOTIC BEHAVIOUR OF THE PROBABILITY-WEIGHTED MOMENTS AND PENULTIMATE APPROXIMATION

JEAN DIEBOLT¹, ARMELLE GUILLOU² AND RYM WORMS³

Abstract. The P.O.T. (Peaks-Over-Threshold) approach consists of using the Generalized Pareto Distribution (GPD) to approximate the distribution of excesses over a threshold. We use the probability-weighted moments to estimate the parameters of the approximating distribution. We study the asymptotic behaviour of these estimators (in particular their asymptotic bias) and also the functional bias of the GPD as an estimate of the distribution function of the excesses. We adapt penultimate approximation results to the case where parameters are estimated.

Mathematics Subject Classification. 60G70, 62G20.


1. INTRODUCTION

Nowadays, it seems common practice to use statistics to describe or even predict the average behaviour of many different phenomena. However, that average behaviour is not the only interesting subject of research in many situations. Indeed, in many problems of considerable interest, it is necessary to estimate the probability of occurrence of extreme values. For instance, one is often interested in the probability that the maximum of $n$ random variables exceeds a given threshold or, vice versa, one wants to determine a level such that the exceedance probability is below a given small value. For example, a hydraulics engineer has to estimate the necessary height of a dike such that the probability of a flooding in a given year is less than $10^{-4}$ (cf. [4]). More precisely, extreme value theory discusses the conditions under which extremes from a random sample $X_1, \ldots, X_n$ are attracted to a non-degenerate limit distribution when the sample size $n$ increases to infinity. This condition seems reasonable when a statistical analysis is based solely on the extremes from the given sample. Here we assume that the random variables $X_i$ ($i = 1, \ldots, n$) are independent and identically distributed with common distribution function $F$. We require that the properly centred and normed sample maxima $X_{n,n} = \max\{X_1, \ldots, X_n\}$ converges in distribution to a non-degenerate limit. Gnedenko (see [11]) showed that this limit distribution is necessarily of extreme value type, i.e. for some $\gamma \in \mathbb{R}$ there exists sequences of constants

---

Keywords and phrases. Extreme values, domain of attraction, excesses, Generalized Pareto Distribution, probability-weighted moments, penultimate approximation.

¹ Université de Marne-la-Vallée, Équipe d’Analyse et de Mathématiques Appliquées, 5 boulevard Descartes, bâtiment Copernic, Champs-sur-Marne, 77454 Marne-la-Vallée Cedex 2, France; e-mail: diebolt@math.univ-mlv.fr
² Université Paris VI, Laboratoire de Statistique Théorique et Appliquée, Bôîte 158, 175 rue du Chevaleret, 75013 Paris, France; e-mail: guillou@ccr.jussieu.fr
³ Université de Rouen, Laboratoire de Mathématiques Raphaël Salem, UMR 6085 du CNRS, Site Colbert, UFR Sciences, 76821 Mont-Saint-Aignan Cedex, France; e-mail: rym.worms@univ-rouen.fr

© EDP Sciences, SMAI 2003
\( \sigma_n > 0 \) and \( \alpha_n \in \mathbb{R} \) such that
\[
\lim_{n \to \infty} \mathbb{P} \left( \frac{X_{n,n} - \alpha_n}{\sigma_n} \leq x \right) \longrightarrow H_\gamma(x)
\]  
for all continuity points of the extreme value distribution function \( H_\gamma \), defined as
\[
H_\gamma(x) = \begin{cases} 
\exp \left( - \left( 1 + \gamma x \right)^{-\frac{1}{\gamma}} \right), & \text{for } \gamma \neq 0 \text{ and } 1 + \gamma x > 0 \\
\exp(\exp(-x)), & \text{for } \gamma = 0 \text{ and } x \in \mathbb{R}.
\end{cases}
\]

The distribution function \( F \) is said to belong to the maximum domain of attraction of \( H_\gamma \). The real-valued parameter \( \gamma \) is referred to as the extreme value index (EVI) of \( F \). Most common continuous distribution functions satisfy this weak condition. Distributions with \( \gamma > 0 \) are called heavy-tailed, as their tail \( F \) typically decays slowly as a power function. Examples in this Fréchet class are the Pareto, Burr, Student’s\( t \), lognormal, gamma and classical Weibull distributions, the tail of which diminishes exponentially fast. Finally, the Weibull class consists of distributions with \( \gamma < 0 \), which all have a finite right endpoint \( s_+(F) := \sup\{x : F(x) < 1\} \). Examples in this class are the uniform and reverse Burr distributions.

The problem of estimating the so-called extreme value index \( \gamma \), which determines the behaviour of the underlying d.f. \( F \) in its upper tail, has received much attention in the last twenty years. An extensive motivation of this estimation problem can be found in [10]. More recently, de Haan and Rootzén (see [15]) demonstrated how to use an estimate of \( \gamma \) to construct estimators of extreme quantiles. Nevertheless, the generalized extreme value distribution is appropriate when the data consist of a set of maxima. However, there has been some criticism of this approach, because using only maxima leads to the loss of information contained in other large sample values in a given period. This problem is remedied by considering several of the largest order statistics instead of just the largest one, that is considering all values larger than a given threshold. The differences between these values and a given threshold are called excesses over the threshold. Denote by \( F_u(x) := \mathbb{P}(X - u \leq x | X > u) \) the distribution of the excesses of \( X \) over \( u \), given that \( u \) is exceeded, and with \( G_{\gamma,\sigma} \) the Generalized Pareto Distribution (GPD) defined, for all \( x \geq 0 \), by
\[
G_{\gamma,\sigma}(x) = \begin{cases} 
1 - \left( \frac{\gamma x}{\sigma} \right)^{-\frac{1}{\gamma}}, & \text{for } \gamma \neq 0 \text{ and } 1 + \gamma x/\sigma > 0 \\
1 - \exp \left( -\frac{x}{\sigma} \right), & \text{for } \gamma = 0.
\end{cases}
\]

Pickands’ and Balkema and de Haan’s result (see [17] and [1]) on the limiting distribution of excesses over a high threshold states that condition (1) holds if and only if
\[
\lim_{u \to s_+(F)} \sup_{0 < x < s_+(F) - u} \left| F_u(x) - G_{\gamma,\sigma(u)}(x) \right| = 0
\]
for some positive scaling function \( \sigma(u) \) depending on \( u \).

Thus, if, for some \( u_n \), one fixes a high threshold \( u_n \) and selects from a sample \( X_1, ..., X_n \) only those observations \( X_{i_1}, ..., X_{i_\alpha} \) that exceed \( u_n \), a GPD with parameters \( \gamma \) and \( \sigma_n = \sigma(u_n) \) is likely to be a good approximation for the distribution \( F_n \) of the \( N_n \) excesses \( Y_j := X_{i_j} - u_n, j = 1, ..., N_n \). This is called the Peaks-Over-Threshold (P.O.T.) method.

As for the generalized extreme value (GEV) distribution, quite a number of techniques have been proposed to fit a GPD \( G_{\gamma,\sigma} \) to the excesses \( Y_1, ..., Y_{N_n} \). Maximum Likelihood Estimation (MLE, see [19]) of the Generalized Pareto Distribution has previously been considered in the literature, but Hosking and Wallis (see [16]) showed, using computer simulation, that, Maximum Likelihood Estimation, although asymptotically the most efficient method, does not clearly display its efficiency even in samples as large as 500, and that, in such cases, estimators derived by the method of probability-weighted moments are more reliable. This motivates in particular the fact
that, in this paper, we use this last mentioned method (PWM) in order to estimate the parameters $\gamma$ and $\sigma_n$. Nevertheless, note that, in a forthcoming paper, we will derive the more general conditions on the estimators under which our results hold. These PWM estimators, denoted by $\hat{\gamma}_n$ and $\hat{\sigma}_n$ (see Sect. 2.1 for definitions), are based on the empirical distribution function of the excesses

$$ k_{n,u_n}(x) = \frac{1}{N_n} \sum_{j=1}^{N_n} I(y_j \leq x). $$

It is of course very important to measure the error between $\bar{F}_{u_n} := 1 - F_{u_n}$ (unknown) and its estimator $\hat{G}_{\gamma_n,\sigma_n} := 1 - \hat{G}_{\gamma_n,\sigma_n}$. This error can be splitted into two parts: an approximation error and an estimation error. The first one, also called bias of approximation, is justified by the fact that the distribution of the excesses over $u_n$ is only approximated by a GPD, which implies a systematic error studied in [20]. Since the distribution of the excesses over $u_n$ is $F_{u_n}$ and not $G_{\gamma,\sigma_n}$, the error due to the estimation of $(\gamma, \sigma_n)$ is also divided into an approximation error due to the bias of approximation and a random term due to fluctuations.

Note that $N_n$ follows a binomial distribution $B(n, 1 - F(u_n))$. We suppose, in all the sequel that, $n(1 - F(u_n)) \to \infty$ as $n \to \infty$, that is $N_n \to \infty$ in probability. In such a case, $\frac{N_n}{n(1 - F(u_n))} \to 1$ in probability as $n \to \infty$.

Let

$$ F_{\sigma_n}(y) = \bar{F}_{u_n}(\sigma_n, y) \quad \text{and} \quad k_{\sigma_n,u_n}(y) = k_{u_n,u_n}(\sigma_n, y) $$

for all $y \in \mathbb{R}_+$. We will study the asymptotic behaviour of $\hat{\gamma}_n$ and $\hat{\sigma}_n$ and more specifically the difference

$$ \bar{F}_{u_n}(x) - \hat{G}_{\gamma_n,\sigma_n}(x) $$

where $\hat{\gamma}_n$ and $\hat{\sigma}_n$ are the PWM estimators introduced by Hosking and Wallis (in [16]). Note that the main difference with Drees (see [7]) is that we consider the estimation of the two parameters $\gamma$ and $\sigma_n$ simultaneously.

In what follows, we suppose that $F$ is twice differentiable and that its inverse $F^{-1}$ exists. Let $V$ and $A$ be two functions defined as

$$ V(t) = F^{-1}(e^{-t}) \quad \text{and} \quad A(t) = \frac{V''(\ln t)}{V'(\ln t)} - \gamma. $$

We suppose the following first and second order conditions:

$$ \lim_{t \to +\infty} A(t) = 0, \quad \lim_{t \to +\infty} A'(t) = 0. $$

and

$$ A \quad \text{is of constant sign at } \infty \quad \text{and there exists } \rho \leq 0 \quad \text{such that } |A| \in RV_{\rho}. $$

Under these assumptions, it is proved in [20] (Th. 1.4, p. 43) that as $u_n \to s_+(F)$

$$ \bar{F}_{u_n}(\sigma_n y) - \hat{G}_{\gamma_n}(y) = a_n D_{\gamma, \rho}(y) + o(a_n), \quad \text{as } n \to +\infty, $$

for all $y$, when

$$ \sigma_n := \sigma(u_n) = V'(V^{-1}(u_n)), \quad a_n := A\left(e^{V^{-1}(u_n)}\right), $$

and

$$ D_{\gamma, \rho}(y) := \begin{cases} C_{0, \rho}(y), & \text{if } \gamma = 0, \\ C_{\gamma, \rho}\left(\frac{1}{\gamma} \ln(1 + \gamma y)\right), & \text{if } \gamma \neq 0, \end{cases} $$

where

$$ C_{\gamma, \rho}(y) := e^{-(1+\gamma)y} I_{\gamma, \rho}(y) \quad \text{and} \quad I_{\gamma, \rho}(y) := \int_0^y e^{\gamma u} \int_0^u e^{\rho s} ds du. $$
We assume also that
\[ \lim_{n \to \infty} \sqrt{n} \left(1 - F(u_n)\right) \, a_n = \lambda \in \mathbb{R}. \] (5)

This is equivalent to suppose that \( \sqrt{N_n a_n} \) tends to \( \lambda \) in probability, as \( n \to \infty \).

The main result of this paper is the following. When \( \rho \) is equal to 0, the error due to the fact that \( \hat{F}_u \) is replaced by \( \hat{F}_{\hat{\gamma}_n, \hat{\sigma}_n} \) is of smaller order than the same error in the case \( \rho \neq 0 \).

This result is closely linked (see Sect. 2.5 for further comments) to the penultimate approximation for the distribution of the excesses established in [21] (Gomes and de Haan in [13], generalizing [3, 12] and [14], studied penultimate approximation for the distribution of the maximum). At first sight, it can appear a bit strange since it is well known that, if we consider only the problem of the estimation of the index \( \gamma \), the smaller \( |\rho| \), the more difficult it is to estimate the index. This problem of bias in the estimation of the index has been widely studied recently in the literature and justified in particular the work on regression model by Feuerverger and Hall (see [9]) and Beirlant et al. (see [2]). This paper proves that, on the contrary, if we consider the problem of the estimation of the tail distribution, we do not need to construct asymptotically unbiased estimators, which is essential in the other estimation problem.

The remainder of our paper is organized as follows. In Section 2, we introduce the PWM estimators of Hosking and Wallis and study their asymptotic properties, from which we derive the asymptotic normality of \( (\hat{\gamma}_n, \hat{\sigma}_n) \). Then, we establish our functional result and adapt penultimate approximation results to the case where the parameters are estimated by the PWM estimators. In Section 3, we provide some simulated examples in order to illustrate our main result. Finally, in Section 4, we describe the formalism used. The details of the proofs are postponed to the Appendix.

2. Probability-weighted moments estimators

2.1. Definitions

Define, as in [16], the two first probability weighted moments of the GPD by,
\[ \mu_0(\gamma, \sigma) = \int_0^\infty x \, dG_{\gamma, \sigma}(x) \quad \text{and} \quad \mu_1(\gamma, \sigma) = \int_0^\infty x \, (1 - G_{\gamma, \sigma}(x)) \, dG_{\gamma, \sigma}(x). \] (6)

Note that
\[ \mu_0(\gamma, \sigma) = \int_0^\infty (1 - G_{\gamma, \sigma}(x)) \, dx \quad \text{and} \quad \mu_1(\gamma, \sigma) = \int_0^\infty \frac{1}{2} \, (1 - G_{\gamma, \sigma}(x))^2 \, dx. \] (7)

When \( \gamma < 1 \),
\[ \mu_0(\gamma, \sigma) = \frac{\sigma}{1 - \gamma} \quad \text{and} \quad \mu_1(\gamma, \sigma) = \frac{\sigma}{2(2 - \gamma)}, \] (8)

so that
\[ \gamma = 2 - \frac{\mu_0(\gamma, \sigma)}{\mu_0(\gamma, \sigma) - 2\mu_1(\gamma, \sigma)} \quad \text{and} \quad \sigma = \frac{2\mu_0(\gamma, \sigma)\mu_1(\gamma, \sigma)}{\mu_0(\gamma, \sigma) - 2\mu_1(\gamma, \sigma)}. \] (9)

If we now replace \( \mu_0(\gamma, \sigma) \) and \( \mu_1(\gamma, \sigma) \) by their empirical moment estimators, one obtains the probability-weighted moments estimators \( \hat{\gamma}_n \) and \( \hat{\sigma}_n \) defined as:
\[ \hat{\gamma}_n = 2 - \frac{\hat{\mu}_{0,n}}{\hat{\mu}_{0,n} - 2\hat{\mu}_{1,n}} \quad \text{and} \quad \hat{\sigma}_n = \frac{2\hat{\mu}_{0,n}\hat{\mu}_{1,n}}{\hat{\mu}_{0,n} - 2\hat{\mu}_{1,n}}, \] (10)
where
\[ \hat{\mu}_{0,n} = \int_0^\infty x \, dA_{n,u_n}(x) = \sigma_n \int_0^\infty y \, dA_{n,u_n}^*(y) = \sigma_n \int_0^\infty [1 - A_{n,u_n}^*(y)] \, dy \quad (11) \]
and
\[ \hat{\mu}_{1,n} = \int_0^\infty \frac{1}{2} [1 - A_{n,u_n}(x)]^2 \, dx = \sigma_n \int_0^\infty \frac{1}{2} [1 - A_{n,u_n}^*(y)]^2 \, dy. \quad (12) \]

As mentioned in the Introduction, Hosking and Wallis give formulae for the approximate standard errors of these estimators. They compare their approach to the MLE approach and come to the conclusion that in case \( \gamma \geq 0 \) the method of probability-weighted moments offers a viable alternative. In the sequel, we note \( \mu_0 = \mu_0(\gamma, 1) \) and \( \mu_1 = \mu_1(\gamma, 1) \).

We are going to study the asymptotic behaviour of \((\hat{\gamma}_n, \hat{\sigma}_n/\sigma_n)\) as \( n \to +\infty \). Note that Drees (in [7]) has already studied the asymptotic behaviour of \( \hat{\gamma}_n \) using a different approach.

\[\text{2.2. Asymptotic behaviour of } \hat{\mu}_{0,n} \text{ and } \hat{\mu}_{1,n} \]

**Theorem 1.** Under assumptions (2, 3) with \(-1 < \gamma < 1/2\), and
\[ \lim_{n \to \infty} \sqrt{n} (1 - F(u_n)) a_n = \lambda, \lambda \in \mathbb{R}, \text{ as } n \to +\infty, \quad (13) \]
we have, for almost all sequences \( k_n \to +\infty \), conditionally on \( N_n = k_n \),
\[ \sqrt{k_n} \left( \begin{array}{c} \hat{\mu}_{0,n} - \mu_0 \\ \hat{\mu}_{1,n} - \mu_1 \end{array} \right) \xrightarrow{d} \mathcal{N}(\lambda C, \Gamma), \]

where
\[ \Gamma = \left( \begin{array}{cc} \frac{1}{(1 - \gamma)^2(1 - 2\gamma)} & \frac{1}{2(1 - \gamma)^2(2 - \gamma)} \\ \frac{1}{2(1 - \gamma)^2(2 - \gamma)} & \frac{1}{(2 - \gamma)^2(3 - 2\gamma)} \end{array} \right) \quad \text{and} \quad C = \left( \begin{array}{c} 1 \\ 
\frac{1}{(1 - \gamma)(1 - \gamma + |\rho|)} \end{array} \right). \]

Moreover, if \( \lambda = 0 \), then
\[ \left( \begin{array}{c} \hat{\mu}_{0,n} - \mu_0 \\ \hat{\mu}_{1,n} - \mu_1 \end{array} \right) = a_n C + o(a_n) + \frac{1}{\sqrt{k_n}} \left( \begin{array}{c} \xi_n^0 \\ \xi_n^1 \end{array} \right) + O_P \left( \frac{1}{k_n} \right), \]

where \( \left( \xi_n^0, \xi_n^1 \right) \) converges in distribution to a \( \mathcal{N}(0, \Gamma) \) distribution.

**Proof.**
- We need the following lemma which will be proved in Appendix 5.1.
Lemma 1.

(1) We have
\[ \int_0^\infty (\bar{F}_{u_n}^*(y) - \bar{G}_\gamma(y)) \, dy = \frac{a_n}{(1 - \gamma)(1 - \gamma + |\rho|)} + o(a_n). \] (14)

(2) Moreover, we have
\[ \frac{1}{2} \int_0^{+\infty} \left( (\bar{F}_{u_n}^*(y))^2 - (\bar{G}_\gamma(y))^2 \right) \, dy = \frac{a_n}{2(2 - \gamma)(2 - \gamma + |\rho|)} + o(a_n). \] (15)

Write
\[ \hat{\mu}_{0,n,s} - \mu_0 = \int_0^\infty (1 - \bar{K}_{u_n}^*(y)) \, dy - \int_0^\infty (1 - G_\gamma(y)) \, dy. \] (16)

Lemma 1 yields
\[ \int_0^\infty (\bar{F}_{u_n}^*(y) - \bar{G}_\gamma(y)) \, dy = \frac{a_n}{(1 - \gamma)(1 - \gamma + |\rho|)} + o(a_n). \] (19)

In order to study the second term in the right-hand side of (18), we need the following lemma established in Einmahl and Mason (see [8]).

Lemma 2. On a rich enough probability space, there exist a sequence of probabilistically equivalent versions \( \tilde{\alpha}_n(.) \) of the uniform empirical process \( \alpha_n(.) \) and a fixed Brownian bridge \( B \) such that, for all \( \nu \) satisfying \( 0 \leq \nu < 1/4 \), we have
\[ \sup_{1/n \leq t \leq 1 - 1/n} n^\nu \frac{|\tilde{\alpha}_n(t) - B(t)|}{(t(1 - t))^{1/2 - \nu}} = O_P(1). \] (20)

Let \( \tilde{\alpha}_n \) be such a processus for which we can apply (20). We prove in Appendix 5.2 that for \( \gamma < 1/2 \), \( \int_0^\infty \tilde{\alpha}_n (F_{u_n}^*(y)) \, dy \) converges in probability to \( \int_0^1 h_\gamma(t) \, dB(t) \), where
\[ h_\gamma(t) = \begin{cases} \frac{t^{-\gamma} - 1}{\gamma} & \text{if } \gamma \neq 0, \\ \ln(1/t) & \text{if } \gamma = 0. \end{cases} \] (21)

It follows that
\[ \frac{\hat{\mu}_{0,n,s}}{\sigma_n} - \mu_0 = \left( \frac{a_n}{(1 - \gamma)(1 - \gamma + |\rho|)} + o(a_n) \right) + \frac{1}{\sqrt{k_n}} \xi_0. \] (22)
where

$$
\xi_n^0 \xrightarrow{P} - \int_0^1 h_{\gamma}(t) dB(t).
$$

• Now, similarly we have

$$
\hat{\mu}_{1,n} - \mu_1 \overset{d}{=} \int_{0}^{+\infty} \frac{1}{2} \left( 1 - K_{K_n, u_n}(y) \right)^2 \, dy - \int_{0}^{+\infty} \frac{1}{2} \left( 1 - G_{\gamma}(y) \right)^2 \, dy

\overset{d}{=} \frac{1}{2} \int_{0}^{+\infty} \left( (F_{u_n}^*(y))^2 - (G_{\gamma}(y))^2 \right) \, dy - \frac{1}{\sqrt{k_n}} \int_{0}^{+\infty} F_{u_n}^*(y) \alpha_{k_n} \left( F_{u_n}^*(y) \right) \, dy

+ \frac{1}{2k_n} \int_{0}^{+\infty} \left( \alpha_{k_n} \left( F_{u_n}^*(y) \right) \right)^2 \, dy.
$$

We are going to study separately the three terms of the last equality.

Lemma 1 yields

$$
\frac{1}{2} \int_{0}^{+\infty} \left( (F_{u_n}^*(y))^2 - (G_{\gamma}(y))^2 \right) \, dy = \frac{a_n}{2(2 - \gamma)(2 - \gamma + |\rho|)} + o(a_n).
$$

The same proof as in Appendix 5.2, shows that for \( \gamma < 1/2 \), \( \int_0^{+\infty} F_{u_n}^*(y) \bar{\alpha}_{k_n} \left( F_{u_n}^*(y) \right) \, dy \) converges in probability to \( \int_{0}^{1} k_{\gamma}(t) \, dB(t) \), where

$$
k_{\gamma}(t) = - \frac{t^{1-\gamma}}{1 - \gamma}.
$$

We prove in Appendix 5.3 that

$$
\int_{0}^{+\infty} \left( \bar{\alpha}_{k_n} \left( F_{u_n}^*(y) \right) \right)^2 \, dy = O_P \left( \frac{1}{k_n} \right).
$$

Therefore, it follows that

$$
\frac{\hat{\mu}_{1,n}}{\sigma_n} - \mu_1 \overset{d}{=} \frac{a_n}{2(2 - \gamma)(2 - \gamma + |\rho|)} + o(a_n) + \frac{1}{\sqrt{k_n}} \xi_n^1 + O_P \left( \frac{1}{k_n} \right),
$$

where

$$
\xi_n^0 \xrightarrow{P} - \int_0^1 k_{\gamma}(t) \, dB(t).
$$

Finally, as \( \lim_{n \to \infty} \sqrt{n(1 - F(u_n))} \, a_n = \lambda \) and \( N_n/(n(1 - F(u_n))) \xrightarrow{P} 1 \), it follows that, conditionally on \( N_n = k_n \),

$$
\sqrt{k_n} \begin{pmatrix}
\frac{\hat{\mu}_{0,n}}{\sigma_n} - \mu_0 \\
\frac{\hat{\mu}_{1,n}}{\sigma_n} - \mu_1
\end{pmatrix} = \lambda \begin{pmatrix} c_0 \\ c_1 \end{pmatrix} + \left( \frac{\xi_0}{\xi_n} \right) + o_P(1),
$$

where

$$
\xi_n^0 \xrightarrow{P} - \int_0^1 h_{\gamma}(t) dB(t).
$$
where

\[
\begin{pmatrix}
  c_0 \\
  c_1 
\end{pmatrix} = \begin{pmatrix}
  1 \\
  \frac{1}{(1-\gamma)(1-\gamma+|\rho|)} \\
  \frac{1}{2(2-\gamma)(2-\gamma+|\rho|)}
\end{pmatrix}
\]  \tag{27}

and

\[
\begin{pmatrix}
  \xi^0_n \\
  \xi^1_n 
\end{pmatrix} \overset{d}{\longrightarrow} \begin{pmatrix}
  Y_0 \\
  Y_1 
\end{pmatrix} = \begin{pmatrix}
  -\int_0^1 h_\gamma(t) \, dB(t) \\
  -\int_0^1 k_\gamma(t) \, dB(t)
\end{pmatrix}
\]  \tag{28}

which is a \( N(0, \Gamma) \) distribution,

\[
\Gamma = \begin{pmatrix}
  \text{Var}(Y_0) & \text{Cov}(Y_0, Y_1) \\
  \text{Cov}(Y_0, Y_1) & \text{Var}(Y_1)
\end{pmatrix}
\]

\[
= \begin{pmatrix}
  1 & 1 \\
  1 & 2(1-\gamma)^2(2-\gamma)
\end{pmatrix}
\]  \tag{29}

2.3. Asymptotic behaviour of \( \hat{\gamma}_n \) and \( \hat{\sigma}_n \)

**Corollary 1.** Under the same assumptions as Theorem 1, if we suppose moreover that\(^4\)

\[
n(1-F(u_n)) a_n \to \infty \text{ as } n \to +\infty,
\]  \tag{30}

then, for almost all sequences \( k_n \to +\infty \), we have, conditionally on \( N_n = k_n \), that

\[
\sqrt{k_n} \begin{pmatrix}
  \hat{\gamma}_n - \gamma \\
  \hat{\sigma}_n - \sigma
\end{pmatrix} \overset{d}{\longrightarrow} N(\lambda B, \Sigma),
\]

where

\[
\Sigma = \frac{1}{(1-2\gamma)(3-2\gamma)} \begin{pmatrix}
  (1-\gamma)(2-\gamma)^2(1-\gamma+2\gamma^2) & (2-\gamma)(2-6\gamma + 7\gamma^2 - 2\gamma^3) \\
  (2-\gamma)(2-6\gamma + 7\gamma^2 - 2\gamma^3) & 7 - 18\gamma + 11\gamma^2 - 2\gamma^3
\end{pmatrix}
\]

and

\[
B = \begin{pmatrix}
  \frac{(2-\gamma)(1-\gamma)}{(1-\gamma+|\rho|)(2-\gamma+|\rho|)} \\
  \frac{|\rho|}{(1-\gamma+|\rho|)(2-\gamma+|\rho|)}
\end{pmatrix}
\]  \tag{31}

\(^4\)This condition is trivial in the case \( \lambda \neq 0 \).
Moreover, if $\lambda = 0$, then

$$
\sqrt{k_n} \left( \frac{\hat{\gamma}_n - \gamma}{\hat{\sigma}_n / \sigma_n - 1} \right) - a_n B \overset{d}{\longrightarrow} \mathcal{N}(0, \Sigma).
$$

**Proof.** Define the mapping $\Psi$, from $\mathbb{R}^2 \cap \{(a, b) \in \mathbb{R}^2 : a > 2b\}$ into $]-\infty, 1[\times \mathbb{R}_+$, by

$$
\Psi(a, b) = \left( 2 - \frac{a}{a - 2b}, \frac{2ab}{a - 2b} \right).
$$

Then,

$$
(\gamma, 1) = \Psi (\mu_0, \mu_1)
$$

and

$$
\left( \frac{\hat{\gamma}_n}{\hat{\sigma}_n / \sigma_n} \right) = \Psi \left( \frac{\hat{\mu}_0.n}{\hat{\sigma}_n / \sigma_n}, \frac{\hat{\mu}_1.n}{\hat{\sigma}_n / \sigma_n} \right).
$$

The proof of Theorem 1 implies (see (26, 27) and (28)) that, conditionally on $N_n = k_n$,

$$
\left( \frac{\hat{\mu}_0.n}{\hat{\sigma}_n / \sigma_n}, \frac{\hat{\mu}_1.n}{\hat{\sigma}_n / \sigma_n} \right) \overset{d}{=} \left( \mu_0 / \sigma_n, \mu_1 / \sigma_n \right) + a_n \left( c_0 / c_1 \right) + \frac{1}{\sqrt{k_n}} \left( \xi_0 / \xi_1 \right) + o(a_n) + \mathcal{O}_P \left( \frac{1}{k_n} \right)
$$

(32)

A Taylor expansion yields

$$
\left( \frac{\hat{\gamma}_n - \gamma}{\hat{\sigma}_n / \sigma_n - 1} \right) = D\Psi(\mu_0, \mu_1) \left( \begin{array}{c} \alpha_n \\ \beta_n \end{array} \right) + \frac{1}{2} (\alpha_n \beta_n) D^2\Psi(\mu_0, \mu_1) \left( \begin{array}{c} \alpha_n \\ \beta_n \end{array} \right) + o(a_n^2) + \mathcal{O}_P \left( \frac{1}{k_n} \right),
$$

where

$$
D\Psi(\mu_0, \mu_1) = \frac{2}{(2\mu_1 - \mu_0)^2} \begin{pmatrix} \mu_1 & -\mu_0 \\ -2\mu_1^2 & \mu_0^2 \end{pmatrix}.
$$

(33)

Since $\lim_{n \to \infty} \sqrt{n(1 - F(u_n))} a_n = \lambda$ and $n(1 - F(u_n)) a_n \to \infty$, we have

$$
\frac{1}{2} (\alpha_n \beta_n) D^2\Psi(\mu_0, \mu_1) \left( \begin{array}{c} \alpha_n \\ \beta_n \end{array} \right) = \mathcal{O}_P \left( \frac{1}{k_n} \right).
$$

Moreover,

$$
D\Psi(\mu_0, \mu_1) \left( \begin{array}{c} \alpha_n \\ \beta_n \end{array} \right) = a_n \left( \begin{array}{c} b_0 \\ b_1 \end{array} \right) + \frac{1}{\sqrt{k_n}} D\Psi(\mu_0, \mu_1) \left( \begin{array}{c} \xi_0 \\ \xi_1 \end{array} \right) + o_P(a_n),
$$
where
\[
\begin{pmatrix}
  b_0 \\
  b_1 
\end{pmatrix} = \frac{2}{(2\mu_1 - \mu_0)^2} \begin{pmatrix}
  \mu_1 c_0 - \mu_0 c_1 \\
  -2\mu_1^2 c_0 + \mu_0^2 c_1 
\end{pmatrix}
\]
and
\[
D\Psi(\mu_0, \mu_1) \begin{pmatrix}
  Y_0 \\
  Y_1 
\end{pmatrix}
\]
is a \( \mathcal{N}(0, \Sigma) \) distribution,
with
\[
\Sigma = D\Psi(\mu_0, \mu_1) \Gamma (D\Psi(\mu_0, \mu_1))^T.
\]
Thus,
\[
\sqrt{k_n} \begin{pmatrix}
  \hat{\gamma}_n - \gamma \\
  \hat{\sigma}_n - \sigma_n
\end{pmatrix} \xrightarrow{d} \mathcal{N}(\lambda B, \Sigma),
\]
with \( B = \begin{pmatrix} b_0 \\ b_1 \end{pmatrix} \).

2.4. Functional result

**Theorem 2.** Under the same assumptions as Corollary 1, for almost all sequences \( k_n \to +\infty \), we have, conditionally on \( N_n = k_n \), that

(i) the process
\[
\sqrt{k_n} (F_{u_n}^*(x) - G_{\hat{\gamma}_n, \hat{\sigma}_n}(x))
\]
converges in distribution to
\[
\lambda B(x) + Z(x),
\]
where
\[
B(x) = D_{\gamma, \rho}(x) + L_1(\gamma, \rho) \frac{\partial G_{\gamma}}{\partial \gamma}(x) + L_2(\gamma, \rho) \frac{\partial G_{\gamma}}{\partial x}(x)
\]
with
\[
L_1(\gamma, \rho) = \frac{(2 - \gamma)(1 - \gamma)}{(1 - \gamma + |\rho|)(2 - \gamma + |\rho|)} \quad \text{and} \quad L_2(\gamma, \rho) = -\frac{x|\rho|}{(1 - \gamma + |\rho|)(2 - \gamma + |\rho|)},
\]
and \( Z(x) \) is a centered Gaussian process equal to
\[
\frac{\partial G_{\gamma}}{\partial \gamma}(x) \int_0^1 \ell_1^{(1)}(t) \, dB(t) - x \frac{\partial G_{\gamma}}{\partial x}(x) \int_0^1 \ell_2^{(1)}(t) \, dB(t),
\]
with
\[
\ell_1^{(1)}(t) = \begin{cases} 
  (2 - \gamma)(2\gamma - \gamma t) + (1 - \gamma)^2 \frac{t - \gamma}{\gamma} & \text{if } \gamma \neq 0, \\
  8t - 2\ln t & \text{if } \gamma = 0,
\end{cases}
\]
and
\[
\ell_1^{(2)}(t) = \begin{cases} 
  (2 - \gamma)(2\gamma - \gamma t) + (1 - \gamma)^2 \frac{t - \gamma}{\gamma} & \text{if } \gamma \neq 0, \\
  8t - 2\ln t & \text{if } \gamma = 0,
\end{cases}
\]
and
\[
\ell_2^{(2)}(t) = \begin{cases} 
  (2 - \gamma)(2\gamma - \gamma t) + (1 - \gamma)^2 \frac{t - \gamma}{\gamma} & \text{if } \gamma \neq 0, \\
  8t - 2\ln t & \text{if } \gamma = 0.
\end{cases}
\]
Remark 1.\ B(x) is called the functional approximation error, it contains the approximation error between $F_{u_n}^\ast(x)$ and $G_\gamma(x)$ (see (4)) and the approximation error due to the estimation of the unknown parameters $\gamma$ and $\sigma_n$.

Proof. (i) Write

$$
\bar{F}_{u_n}^\ast(x) - \bar{G}_{\bar{\gamma}_n}(\sigma_n x/\bar{\sigma}_n) = \bar{F}_{u_n}(x) - \bar{G}_\gamma(\sigma_n x/\bar{\sigma}_n) + \bar{G}_\gamma(\sigma_n x/\bar{\sigma}_n) - \bar{G}_{\bar{\gamma}_n}(\sigma_n x/\bar{\sigma}_n),
$$

with

$$
\bar{F}_{u_n}^\ast(x) - \bar{G}_\gamma(\sigma_n x/\bar{\sigma}_n) = \bar{F}_{u_n}(x) - \bar{G}_\gamma(x) + x \left( \frac{\partial G_\gamma}{\partial x} \right) + O_P \left( \frac{1}{k_n} \right) = a_n D_{\gamma,\rho}(x) - x \left( \frac{\partial G_\gamma}{\partial x} \right) + o_P(a_n)
$$

and

$$
\bar{G}_\gamma(\sigma_n x/\bar{\sigma}_n) - \bar{G}_{\bar{\gamma}_n}(\sigma_n x/\bar{\sigma}_n) = (\bar{\gamma}_n - \gamma) \frac{\partial G_\gamma}{\partial \gamma}(\sigma_n x/\bar{\sigma}_n) + O_P \left( \frac{1}{k_n} \right).
$$

The result follows from the facts that, on the first hand (see Cor. 1),

$$
\bar{\gamma}_n - \gamma = a_n b_0 + \frac{\eta_n}{\sqrt{k_n}} + o_P(a_n) \quad \text{and} \quad \bar{\sigma}_n - \sigma_n = 1 = a_n b_1 + \frac{\eta_n}{\sqrt{k_n}} + o_P(a_n),
$$

with

$$
\begin{pmatrix}
  b_0 \\
  b_1
\end{pmatrix}
$$

given by (34) and

$$
\begin{pmatrix}
  \eta_n^0 \\
  \eta_n^1
\end{pmatrix}
\xrightarrow{d}
D\Phi(\mu_0, \mu_1)
\begin{pmatrix}
  Y_0 \\
  Y_1
\end{pmatrix},
$$

and, on the other hand (see (28) and (33)),

$$
D\Phi(\mu_0, \mu_1)
\begin{pmatrix}
  Y_0 \\
  Y_1
\end{pmatrix}
= \frac{2}{(2\mu_1 - \mu_0)^2} \left( \int_0^1 (\mu_1 h_\gamma(t) - \mu_0 k_\gamma(t)) \, d\mathbb{B}(t) \right)
= \left( \int_0^1 \ell_\gamma^{(1)}(t) \, d\mathbb{B}(t) \right) \left( \int_0^1 \ell_\gamma^{(2)}(t) \, d\mathbb{B}(t) \right).
$$
(ii) When $\gamma = 0$, it is easy to check that $D_{\gamma,\rho}(x) = \frac{x^2}{2}e^{-x} - \frac{e^{\gamma x} - 1 - \gamma xe^{\gamma x}}{\gamma^2}$ and hence $C_{\gamma,0}(x) = e^{-(1+\gamma)x}1 + \frac{\gamma x e^{\gamma x} - e^{\gamma x}}{\gamma^2}$. Therefore, as for $\gamma = 0$,

$$D_{\gamma,0}(x) = C_{\gamma,0}\left(\frac{\ln(1+\gamma x)}{\gamma}\right) = -\frac{\partial G_\gamma}{\partial \gamma}(x).$$

(39)

Taking $\rho = 0$ in (37) and using the fact that $D_{\gamma,0}(x) = -\frac{\partial G_\gamma}{\partial \gamma}(x)$, we derive that $B(x) = 0$, which is not the case if $\rho < 0$.

Let us give now some comments concerning this theorem:

- this result means that, in case $\rho = 0$, the error due to the fact that $\bar{F}_{\gamma\alpha}(x)$ is replaced by $G_{\gamma,\bar{\gamma}n/\sigma_n}(x)$ is of smaller order than the error made when one takes $\hat{\gamma}_n$ in order to estimate $\gamma$. As noticed in the Introduction, this result is surprising taking into account the fact that in the case of the estimation of the index $\gamma$, the bias goes up when $\rho$ tends to 0;
- the second order parameter $\rho$ is zero for many usual distributions in the Gumbel domain of attraction ($\gamma = 0$): e.g. the normal, lognormal, gamma and classical Weibull distributions. Hence, our result applies to all these distributions;
- this result is closely linked to penultimate approximation established in [21], which is developed in the next section.

2.5. Penultimate approximation

It is established in [21] (Th. 1) that penultimate approximation is only possible in the case $\rho = 0$. That is, $F_{\gamma n}(x)$ is approximated by $G_{\gamma+\alpha_n}(x)$ rather than by $G_{\gamma}(x)$. When $\rho$ is zero, it is proved in [21] (Th. 3) that under some additional conditions (satisfied by many usual distributions), when $\gamma > -1$,

$$\bar{F}_{\gamma n}(\sigma_n y) - G_{\gamma+\alpha_n}(y) = b_n D_{\gamma}(y) + o(b_n) \text{ as } n \to +\infty,$$  

(40)

where $b_n = o(a_n)$ is defined by

$$b_n = e^{V^{-1}(u_n)}A'(e^{V^{-1}(u_n)})$$  

(41)

and

$$D_{\gamma}(y) = \begin{cases} C_0(y) & \text{if } \gamma = 0, \\ C_\gamma\left(\frac{1}{\gamma}\ln(1+\gamma y)\right) & \text{if } \gamma \neq 0, \end{cases}$$

where

$$C_\gamma(y) = e^{-(1+\gamma)y}M_\gamma(y) \text{ and } M_\gamma(y) = \frac{1}{2}\int_0^yu^2e^{\gamma u}du.$$  

(42)

The same proof as for Corollary 1 shows that the error made by replacing $\gamma + a_n$ by $\hat{\gamma}_n$ is equal to

$$\frac{\gamma^2 - 3\gamma + 3}{2 - \gamma}b_n + o(b_n),$$

where $b_n$ is given by (41).
Now a similar demonstration as the one for Theorem 2 implies that, in the case \( \rho = 0 \), the functional approximation error (analogous of \( B \)), which is of smaller order that \( a_n \), becomes

\[
b_n \left( D_{\gamma}(x) + \frac{\gamma^2 - 3\gamma + 3}{2 - \gamma} \frac{\partial G_{\gamma}}{\partial \gamma}(x) - \frac{\gamma - 1}{2 - \gamma} x \frac{\partial G_{\gamma}}{\partial x}(x) \right) + o(b_n),
\]

which is equal to

\[
b_n \frac{x e^{-x}}{2} \left( 1 - \frac{3x}{2} + \frac{x^2}{2} \right) + o(b_n),
\]

when \( \gamma = 0 \).

**Example.** Let us consider the classical Weibull distribution function defined by

\[
\bar{F}(x) = \exp \left(-x^\beta\right), \quad x \geq 0,
\]

with \( \beta > 0 \). For this distribution, \( \gamma = 0 \) and, when \( \beta \neq 1 \), we have \( V(t) = t^{1/\beta} \). Hence \( A(t) = \left(\frac{1}{\beta} - 1\right) \frac{1}{\ln t} \) and the second order parameter \( \rho \) is equal to 0. This yields

\[
\sigma_n = \frac{1}{\beta} u_n^{1-\beta},
\]

\[
a_n = \left(\frac{1}{\beta} - 1\right) u_n^{-\beta},
\]

\[
b_n = \left(1 - \frac{1}{\beta}\right) u_n^{-2\beta}.
\]

Let \( k_n = n(1 - F(u_n)) \) be the averaged number of excesses over \( u_n \). It follows from the assumption \( \sqrt{n}(1 - F(u_n))a_n \to \lambda \) that \( k_n \) is of order \((\ln n)^2\).

### 3. Simulations

Next we report on a small sample study performed in order to illustrate our theoretical results. With this aim, we consider 100 samples of size \( n = 500 \) drawn from a Burr(\( \beta, \tau, \lambda \)) distribution, defined as

\[
\bar{F}(x):=1-F(x)=\left(\frac{\beta}{\beta+x}\right)^\lambda, \quad x > 0,
\]

where \( \beta, \lambda, \tau > 0 \).

Note that for such a distribution \( \gamma = \frac{1}{\beta} \) and, if \( \tau \neq 1 \), \( \rho = -1/\lambda \). We choose \( k_n = n/10 \) as suggested in [5] and [6] and \((\beta, \tau, \lambda) = (1, \frac{3}{2}, 8); (1, \frac{7}{3}, 10)\) and \((1, \frac{1}{3}, 12)\). For these models \( \gamma = \frac{1}{4} \) and \( \rho = -\frac{1}{8}, -\frac{1}{10} \) and \(-\frac{1}{12} \) respectively.

We take as explicit versions of \( \hat{\mu}_{0,n} \) and \( \hat{\mu}_{1,n} \) the following expressions (see [16] for the choice of \( \hat{\mu}_{1,n} \))

\[
\hat{\mu}_{0,n} = \frac{1}{N_n} \sum_{i=1}^{N_n} Y_i \quad \text{and} \quad \hat{\mu}_{1,n} = \frac{1}{N_n} \sum_{i=1}^{N_n} p_i Y_i,
\]

where \( p_i = 1 - \frac{1-0.35}{N_n} \). For each distribution and each sample we compute the PWM estimators

\[
\hat{\gamma}_n = 2 - \frac{\hat{\mu}_{0,n}}{\hat{\mu}_{0,n} - 2\hat{\mu}_{1,n}} \quad \text{and} \quad \hat{\sigma}_n = \frac{2\hat{\mu}_{0,n}\hat{\gamma}_n}{\hat{\mu}_{0,n} - 2\hat{\mu}_{1,n}}.
\]
for \( j = 1, \ldots, 100 \), and we plot in Figure 1

\[
\left| \tilde{F}_{u_n}(x) - \frac{1}{100} \sum_{j=1}^{100} \tilde{G}_{\hat{\gamma}_j, \hat{\sigma}_j}(x) \right|. 
\]

From this figure it is clear that as \( \rho \) tends to 0, the quantity \( |\tilde{F}_{u_n}(x) - \frac{1}{100} \sum_{j=1}^{100} \tilde{G}_{\hat{\gamma}_j, \hat{\sigma}_j}(x)| \) becomes small. This corroborate the theory obtained in this paper, and proves that if we are interested in the tail distribution, we do not have to use asymptotically unbiased estimators, which is on the other hand essential in the case of the estimation of an extreme value index.

#### 4. Sketch of the Formalism

We denote by \( T(\tilde{F}) \) a class of estimators of \((\gamma, \sigma)\) such that for all \( \gamma \) in some interval and \( \sigma > 0 \),

\[
T(\tilde{G}_{\gamma, \sigma}) = (T_1(\tilde{G}_{\gamma, \sigma}), T_2(\tilde{G}_{\gamma, \sigma})) = (\gamma, \sigma). \tag{43}
\]

We restrict our study to estimators \( T(\tilde{F}) \) satisfying (43) which are scale invariant for all \( \tilde{F} \) in some large class of distributions and all \( \sigma > 0 \),

\[
T_1\left(\tilde{F}\left(\frac{\cdot}{\sigma}\right)\right) = T_1(\tilde{F}) \quad \text{and} \quad T_2\left(\tilde{F}\left(\frac{\cdot}{\sigma}\right)\right) = \sigma T_2(\tilde{F}). \tag{44}
\]

If condition (44) is satisfied, then in particular for all \( \gamma \) and \( \sigma > 0 \),

\[
T_1(\tilde{G}_{\gamma, \sigma}) = T_1(\tilde{G}_{\gamma}) = \gamma \text{ and } T_2(\tilde{G}_{\gamma, \sigma}) = \sigma T_2(\tilde{G}_{\gamma}) = \sigma, \text{ thus } T_2(\tilde{G}_{\gamma}) = 1. \tag{45}
\]

Consequently

\[
\partial T_1(\tilde{G}_{\gamma}) / \partial \gamma = 1 \quad \text{and} \quad \partial T_2(\tilde{G}_{\gamma}) / \partial \gamma = 0. \tag{46}
\]
We suppose moreover a form of Hadamard differentiability: there exist linear forms $DT(G_\gamma)$, defined on a suitable space, such that under assumption (5), we have
\begin{equation}
\lim_{n \to +\infty} \frac{T(\tilde{F}_{u_n}^\ast) - T(G_\gamma)}{a_n} = DT(G_\gamma) [D_{\gamma, \rho}],
\end{equation}
\begin{equation}
\lim_{n \to +\infty} \sqrt{N_n} \left( T \left( \tilde{F}_{u_n}^\ast + \frac{1}{\sqrt{N_n}} \alpha_{N_n} (F_{u_n}^\ast) \right) - T(\tilde{F}_{u_n}^\ast) \right) = DT(G_\gamma) \left[ \mathbb{B} \circ G_\gamma \right],
\end{equation}
in probability, and
\begin{equation}
\lim_{n \to +\infty} \frac{T(G_{\gamma + a_n}) - T(G_\gamma)}{a_n} = DT(G_\gamma) \left[ \frac{\partial G_\gamma}{\partial \gamma} \right].
\end{equation}

We defined $\sigma_n$ as in the Introduction. Thus, denoting by $(\gamma_n^b, \sigma_n^b) = T(\tilde{F}_{u_n})$ the values of $\gamma$ and $\sigma$ which could be obtained by replacing $G_{\gamma, \sigma_n}$ by the excesses distribution, we have
\begin{equation}
\left( \gamma_n^b - \gamma, \frac{\sigma_n^b}{\sigma_n} - 1 \right) = T(G_{\gamma + a_n}) - T(G_\gamma) = a_n DT(G_\gamma) [D_{\gamma, \rho}] + o(a_n)
\end{equation}
\begin{equation}
= a_n \left( \mathcal{L}_1(\gamma, \rho), \mathcal{L}_2(\gamma, \rho) \right) + o(a_n).
\end{equation}

This heuristic reasoning shows that the approximation error due to the estimation of the parameters $\gamma$ and $\sigma_n$ is of order $a_n$. We can now infer, always heuristically, the principal term in the asymptotic expansion (in terms of $a_n$) of $\tilde{F}_{u_n}^\ast - G_{\gamma_n^b, \sigma_n^b/\sigma_n}^\ast$, by calculating the first order approximation between $G_{\gamma_n^b, \sigma_n^b/\sigma_n}$ and $G_{\gamma,1}$. This principal term is the function denoted by $B$ in Section 2.4 and called the functional approximation error. It contains not only the approximation error between $\tilde{F}_{u_n}^\ast(x)$ and $G_\gamma(x)$ (see (47)) but also the approximation error due to the estimation of the unknown parameters $\gamma$ and $\sigma_n$ (see (50)). It is given by
\begin{equation}
B(x) = D_{\gamma, \rho}(x) + \mathcal{L}_1(\gamma, \rho) \frac{\partial G_\gamma}{\partial \gamma}(x) - \mathcal{L}_2(\gamma, \rho) x \frac{\partial G_\gamma}{\partial x}(x).
\end{equation}
In the special case $\rho = 0$, we find that $\mathcal{L}_1(\gamma, 0) = 1$ and $\mathcal{L}_2(\gamma, 0) = 0$ which imply that the function $B$ is equal to zero.

5. Appendix

5.1. Proof of Lemma 1

5.1.1.

Our aim is to prove that
\begin{equation}
\int_0^\infty \frac{\tilde{F}_{u_n}^\ast(y) - G_\gamma(y)}{a_n} \, dy \text{ converges to } \frac{1}{(1-\gamma)(1-\gamma + |\rho|)}, \text{ as } n \to \infty.
\end{equation}

We only treat the case\footnote{The cases $\gamma = 0$ and $\gamma < 0$ are similar.} $\gamma > 0$.

Since
\[ G_\gamma(y) = G_0 \left( \frac{1}{\gamma} \ln (1 + \gamma y) \right) \]
and
\[ \tilde{F}_{u_n}^\ast(y) = G_0 \left( V^{-1}(u_n + \sigma_n y) - V^{-1}(u_n) \right), \]

\[ dG_\gamma(x) = \frac{1}{\gamma} \frac{d}{dx} \ln (1 + \gamma x). \]
We are going now to prove that

\[ \int_0^\infty \frac{\bar{F}_{u_n}^*(y) - \bar{G}_\gamma(y)}{a_n} \, dy = \frac{1}{a_n} \int_0^\infty \exp \left( V^{-1}(u_n) - V^{-1}(u_n + \sigma_n y) \right) \, dy \]

\[ - \frac{1}{a_n} \int_0^\infty \exp \left( -\frac{1}{\gamma} \ln (1 + \gamma y) \right) \, dy. \]

Again we only treat the case

\[ \gamma > 0. \]

Write \( \gamma = \frac{1}{\epsilon} \). For sufficiently small \( \epsilon > 0 \), the right-hand side of (52) becomes

\[ \int_0^\infty \frac{\bar{F}_{u_n}^*(y) - \bar{G}_\gamma(y)}{a_n} \, dy = \frac{1}{a_n} \int_0^\infty e^{-s} \frac{V'(s + V^{-1}(u_n))}{V'(V^{-1}(u_n))} - e^{\gamma s} \, ds. \]

Taking \( s \) equal to \( V^{-1}(u_n + \sigma_n y) - V^{-1}(u_n) \) in the first (resp. \( (1 + \gamma y)/\gamma \) in the second) term of the right-hand side of the preceding formula, we obtain

\[ \int_0^\infty \frac{\bar{F}_{u_n}^*(y) - \bar{G}_\gamma(y)}{a_n} \, dy = \frac{1}{a_n} \int_0^\infty e^{-s} \frac{V'(s + V^{-1}(u_n))}{V'(V^{-1}(u_n))} - e^{\gamma s} \, ds. \]

It is well-known (see (1.22) in [20]) that under (2) and (3),

\[ \lim_{n \to +\infty} \frac{1}{a_n} \left( \frac{V'(s + V^{-1}(u_n))}{V'(V^{-1}(u_n))} - e^{\gamma s} \right) = e^{\gamma s} \int_0^s e^{\rho t} \, dt, \]

and for all \( \epsilon > 0 \) and \( n \) sufficiently large,

\[ e^{-s} \frac{V'(s + V^{-1}(u_n))}{V'(V^{-1}(u_n))} - e^{\gamma s} \leq C e^{(\gamma + \epsilon - 1)s} \int_0^s e^{\rho t} \, dt, \]

for some constant \( C > 0 \). \( (53) \)

Since \( \rho \leq 0 \), the right-hand side of (53) is integrable for \( \epsilon \) sufficiently small and \( \gamma < 1/2 \). Then the result follows by applying Lebesgue dominated convergence.

5.1.2.

Since \( \lim_{y \to +\infty} y \left( \bar{F}_{u_n}^*(y) \right)^2 = 0 \), for \( \gamma < \frac{1}{2} \), it follows by integration by parts that the quantity of interest can be written as

\[ \frac{1}{2} \int_0^\infty \left( (\bar{F}_{u_n}^*(y))^2 - (\bar{G}_\gamma(y))^2 \right) \, dy = \int_0^\infty y \bar{F}_{u_n}^*(y) \, d\bar{F}_{u_n}^*(y) - \int_0^\infty y \bar{G}_\gamma(y) \, d\bar{G}_\gamma(y). \]

We are going now to prove that

\[ \frac{1}{a_n} \left( \int_0^\infty y \bar{F}_{u_n}^*(y) \, d\bar{F}_{u_n}^*(y) - \int_0^\infty y \bar{G}_\gamma(y) \, d\bar{G}_\gamma(y) \right) \]

converges to \( \frac{1}{2(2 - \gamma)(2 - \gamma + \rho)} \), as \( n \to \infty \).

Again we only treat the case \( \gamma > 0 \). Write

\[ \int_0^\infty y \bar{F}_{u_n}^*(y) \, d\bar{F}_{u_n}^*(y) - \int_0^\infty y \bar{G}_\gamma(y) \, d\bar{G}_\gamma(y) = \int_0^\infty y \left( \bar{G}_\gamma(y) - \bar{F}_{u_n}^*(y) \right) \, d\bar{G}_\gamma(y) \]

\[ + \int_0^\infty y \bar{F}_{u_n}^*(y) \, d \left( \bar{G}_\gamma(y) - \bar{F}_{u_n}^*(y) \right). \]

We study separately the two terms of the right-hand side.
We study separately the two terms of the right-hand side of (54).

\[
\int_0^\infty y \left( G_\gamma(y) - F_{u_n}^*(y) \right) dG_\gamma(y) = \int_0^\infty e^{-s} \frac{V(s + V^{-1}(u_n)) - u_n}{V'(V^{-1}(u_n))} dB_{u_n}(s) \\
+ \int_0^\infty e^{-2s} \left( \frac{V(s + V^{-1}(u_n)) - u_n}{V'(V^{-1}(u_n))} - \frac{e^{\gamma s} - 1}{\gamma} \right) ds,
\]

where

\[
B_{u_n}(s) = \tilde{G}_0(s) - \tilde{G}_0 \left( \frac{1}{\gamma} \ln \left( 1 + \frac{V(s + V^{-1}(u_n)) - u_n}{V'(V^{-1}(u_n))} \right) \right),
\]

and similarly

\[
\int_0^\infty y F_{u_n}^*(y) d \left( G_\gamma(y) - F_{u_n}^*(y) \right) = - \int_0^\infty e^{-s} \frac{V(s + V^{-1}(u_n)) - u_n}{V'(V^{-1}(u_n))} dB_{u_n}(s).
\]

Consequently it follows that

\[
\frac{1}{a_n} \left( \int_0^\infty y F_{u_n}^*(y) dF_{u_n}^*(y) - \int_0^\infty y G_\gamma(y) dG_\gamma(y) \right) = \int_0^\infty e^{-2s} \frac{V(s + V^{-1}(u_n)) - u_n}{V'(V^{-1}(u_n))} dB_{u_n}(s).
\]

It is well-known (see (1.23) in [20]) that under (2) and (3),

\[
\lim_{n \to +\infty} \frac{1}{a_n} \left( \frac{V(s + V^{-1}(u_n)) - u_n}{V'(V^{-1}(u_n))} - \frac{e^{\gamma s} - 1}{\gamma} \right) = \int_0^s e^{\gamma t} \int_0^z e^t dt dz,
\]

and, for all \( \epsilon > 0 \) and \( n \) sufficiently large,

\[
\frac{e^{-2s}}{a_n} \left( \frac{V(s + V^{-1}(u_n)) - u_n}{V'(V^{-1}(u_n))} - \frac{e^{\gamma s} - 1}{\gamma} \right) \leq C e^{(\epsilon-2)s} \int_0^s e^{\gamma t} \int_0^z e^t dt dz.
\]

The result now follows, once again, by applying Lebesgue dominated convergence.

5.2. **Proof that** \( \int_0^\infty \tilde{\alpha}_{k_n} \left( F_{u_n}^*(y) \right) dy \) **converges to** \( \int_0^1 h_\gamma(t) d\mathbb{E}(t) \)

Write \( F_{u_n}^*(y) = \tilde{G}_0 \left( V^{-1}(u_n + \sigma_n y) - V^{-1}(u_n) \right) \) and take \( s \) equal to \( V^{-1}(u_n + \sigma_n y) - V^{-1}(u_n) \) in order to obtain that

\[
\int_0^\infty \tilde{\alpha}_{k_n} \left( F_{u_n}^*(y) \right) dy = \int_0^\infty \tilde{\alpha}_{k_n} \left( 1 - e^{-s} \right) \frac{V'(s + V^{-1}(u_n))}{V'(V^{-1}(u_n))} ds \\
= \int_0^\infty \tilde{\alpha}_{k_n} \left( 1 - e^{-s} \right) \left( \frac{V'(s + V^{-1}(u_n))}{V'(V^{-1}(u_n))} - e^{\gamma s} \right) ds + \int_0^\infty \tilde{\alpha}_{k_n} \left( 1 - e^{-s} \right) e^{\gamma s} ds. \tag{54}
\]

We study separately the two terms of the right-hand side of (54).

- **Note that**

\[
\int_0^\infty \tilde{\alpha}_{k_n} \left( 1 - e^{-s} \right) e^{\gamma s} ds = \int_0^1 \tilde{\alpha}_{k_n} (1 - t) \frac{dt}{1 + \gamma}. \tag{55}
\]
is well-defined. The right-hand side of (55) can be split into three terms as follows:

\[
\int_0^{1/k_n} \tilde{\alpha}_{k_n} (1 - t) \frac{dt}{t^{1+\gamma}} = \int_0^{1/k_n} \tilde{\alpha}_{k_n} (1 - t) \frac{dt}{t^{1+\gamma}} + \int_{1/k_n}^{1-1/k_n} \tilde{\alpha}_{k_n} (1 - t) \frac{dt}{t^{1+\gamma}} + \int_{1-1/k_n}^1 \tilde{\alpha}_{k_n} (1 - t) \frac{dt}{t^{1+\gamma}}.
\]

(56)

- Let \( M > \sqrt{2} \) and \( \epsilon > 0 \). It follows from Jaeschke theorem (see Th. 1, p. 600, Chap. 16 in [18]) that there exists \( n_0 \) such that for all \( n \geq n_0 \) and \( t \in [0, 1] \),

\[
\tilde{\alpha}_{k_n} (1 - t) \leq M \sqrt{\ln \ln k_n \ln(1 - t)}, \text{ with probability } 1 - \epsilon.
\]

Then we derive that

\[
\int_0^{1/k_n} \tilde{\alpha}_{k_n} (1 - t) \frac{dt}{t^{1+\gamma}} \leq M \sqrt{\ln \ln k_n} \int_0^{1/k_n} \sqrt{t(1 - t)} \frac{dt}{t^{1+\gamma}}, \text{ with probability } 1 - \epsilon.
\]

Since

\[
\int_0^{1/k_n} \sqrt{t(1 - t)} \frac{dt}{t^{1+\gamma}} \leq \int_0^{1/k_n} \sqrt{t} \frac{dt}{t^{1+\gamma}}
\]

and, for \( \gamma < \frac{1}{2} \),

\[
\sqrt{\ln \ln k_n} \int_0^{1/k_n} \sqrt{t} \frac{dt}{t^{1+\gamma}} = \sqrt{\ln \ln k_n} k_n^{\frac{1}{2} - \frac{1}{2}} \to 0 \text{ as } k_n \to +\infty,
\]

we have

\[
\int_0^{1/k_n} \tilde{\alpha}_{k_n} (1 - t) \frac{dt}{t^{1+\gamma}} \to 0 \text{ as } k_n \to +\infty.
\]

- Similarly, we can prove that

\[
\int_{1-1/k_n}^1 \tilde{\alpha}_{k_n} (1 - t) \frac{dt}{t^{1+\gamma}} \to 0 \text{ as } k_n \to +\infty.
\]

- Therefore, it remains to study the term

\[
\int_{1/k_n}^{1-1/k_n} \tilde{\alpha}_{k_n} (1 - t) \frac{dt}{t^{1+\gamma}}.
\]

Let \( \nu < \frac{1}{2} \), \( M > 0 \) and \( \epsilon > 0 \). Then, it follows from (20) that there exists \( n_1 \) such that for all \( n \geq n_1 \) and \( t \in [\frac{1}{k_n}, 1 - \frac{1}{k_n}] \),

\[
\tilde{\alpha}_{k_n} (1 - t) \leq \mathbb{B}(1 - t) + Mk_n^{-\nu}(t(1 - t))^{1/2-\nu}, \text{ with probability } 1 - \epsilon.
\]
Hence,
\[
\int_{1/k_n}^{1-1/k_n} \tilde{\alpha}_{k_n}(1-t) \frac{dt}{t^{1+\gamma}} \leq \int_{1/k_n}^{1-1/k_n} B(1-t) \frac{dt}{t^{1+\gamma}} + M k_n^{-\nu} \int_{1/k_n}^{1-1/k_n} (t(1-t))^{1/2-\nu} \frac{dt}{t^{1+\gamma}},
\]
with probability > 1 - \epsilon.

Since \( \gamma < \frac{1}{2} \), the two integrals in the right-hand side converge according to the law of iterated logarithm for the Brownian motion.

Finally
\[
\lim_{n \to \infty} \int_{1/k_n}^{1-1/k_n} \tilde{\alpha}_{k_n}(1-t) \frac{dt}{t^{1+\gamma}} \leq \int_0^1 B(1-t) \frac{dt}{t^{1+\gamma}} + a_P(1).
\]

The same computations can be applied in order to obtain the lower bound. Consequently
\[
\lim_{n \to \infty} \int_{1/k_n}^{1-1/k_n} \tilde{\alpha}_{k_n}(1-t) \frac{dt}{t^{1+\gamma}} = \int_0^1 B(1-t) \frac{dt}{t^{1+\gamma}} = \int_0^1 h_\gamma(t) dt.
\]

- It remains to prove that \( \int_0^\infty \tilde{\alpha}_{k_n}(1-e^{-s})(V'(s + V^{-1}(u_n))/V'(V^{-1}(u_n)) - e^{\gamma s}) ds \) tends to 0 in probability, as \( n \) tends to +\( \infty \). We know that, under assumptions (2) and (3) (see Prop. 1.1 in [20]), for \( \epsilon > 0 \) and \( n \) sufficiently large
\[
\frac{V'(s + V^{-1}(u_n))}{V'(V^{-1}(u_n))} - e^{\gamma s} \leq C |a_n| e^{(\gamma+\epsilon)s} \int_0^s e^{\epsilon t} dt.
\]

Moreover, using Theorems 1 and 3 (p. 604 of [18]), there exist \( M > 0 \) and \( n_0(\omega) \in \mathbb{N} \) almost surely finite such that for \( n \geq n_0, s > 0 \) and \( \delta > 0 \)
\[
|\tilde{\alpha}_{k_n}(1-e^{-s})| \leq M e^{-s/2}(\ln k_n)^{1/2+\delta} \quad \text{a.s.}
\]

It follows from (58) and (59) that
\[
\int_0^\infty |\tilde{\alpha}_{k_n}(1-e^{-s})| \left| \frac{V'(s + V^{-1}(u_n))}{V'(V^{-1}(u_n))} - e^{\gamma s} \right| ds \leq C |a_n| (\ln k_n)^{1/2+\delta} \int_0^\infty e^{(\gamma+\epsilon-\frac{1}{2})s} \int_0^s e^{\epsilon t} dt ds.
\]

Since \( \gamma < 1/2 \), one can choose \( \epsilon \) sufficiently small in order to have \( \gamma + \epsilon - 1/2 < 0 \). Moreover, \( a_n(\ln k_n)^{1/2+\delta} \) converges to 0 as \( n \) tends to +\( \infty \) under the assumptions.

Hence \( \int_0^\infty \alpha_{k_n}(1-e^{-s})(V'(s + V^{-1}(u_n))/V'(V^{-1}(u_n)) - e^{\gamma s}) ds \) converges to 0 in probability.

5.3. **Proof that** \( \frac{1}{k_n} \int_{0}^{+\infty} (\tilde{\alpha}_{k_n}(F_{u_n}^*)(y))^2 dy = \mathcal{O}_P \left( \frac{1}{k_n} \right) \)

Using the same methodology as in 5.2, we have
\[
\int_0^{+\infty} (\tilde{\alpha}_{k_n}(F_{u_n}^*)(y))^2 dy = \int_0^{+\infty} (\tilde{\alpha}_{k_n}(1-e^{-s}))^2 \left( \frac{V'(s + V^{-1}(u_n))}{V'(V^{-1}(u_n))} - e^{\gamma s} \right) ds \nonumber
\]
\[+ \int_0^{+\infty} (\tilde{\alpha}_{k_n}(1-e^{-s}))^2 e^{\gamma s} ds,
\]
with the first integral converging to zero in probability. Therefore we have to establish that
\[
\int_0^{+\infty} (\tilde{\alpha}_{k_n}(1-e^{-s}))^2 e^{\gamma s} ds = \mathcal{O}_P(1).
\]

\[\text{ASYMPTOTIC BEHAVIOUR OF THE PWM ESTIMATORS \quad 237}\]
Again this integral can be splitted into three terms. Following exactly the same lines as in 5.2 we can easily prove that these terms tend to zero in probability, which achieves the proof.

**References**


