

Consistency Bands for the Mean Excess Function and Application to Graphical Goodness-of-fit Test for Financial Data

Amadou Diadie Ba¹, El Hadji Deme¹, Cheikh Tidiane Seck^{1,3}, Gane Samb Lo^{1,2}

¹ LERSTAD, Université Gaston Berger, Sénégal

² LSA, Université Pierre et Marie Curie, France.

³ Université Alioune Diop de Bambey, Sénégal.

Correspondence : Cheikh Tidiane Seck, Université Alioune Diop de Bambey, Sénégal. Tel: 221-77-489-6246. E-mail: cheikhtidiane.seck@uadb.edu.sn

Received: December 4, 2015 Accepted: January 7, 2016 Online Published: January 25, 2016

doi:10.5539/jmr.v8n1p42 URL: <http://dx.doi.org/10.5539/jmr.v8n1p42>

Abstract

In this paper, we use the modern setting of functional empirical processes and recent techniques on uniform estimation for non parametric objects to derive consistency bands for the mean excess function in the i.i.d. case. We apply our results for modelling Dow Jones data to see how good the Generalized hyperbolic distribution fits monthly data.

Keywords: consistency bands, Donsker classes, functional empirical processes, generalized hyperbolic distributions, mean excess function, Vapnik-Chervonenkis classes

1. Introduction

Because of its growing importance, the mean excess function, also known as the mean residual life function, has been extensively studied in the literature. It is one of the most useful tools in survival analysis, actuarial science, economics, engineering, and environmental science. For example, the mean excess function plot can serve for fitting Generalized Pareto distributions if it is linear. It is defined as follows.

Let X be a random variable defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, and let F be its distribution function with right-endpoint

$$x_F = \sup\{x \in \mathbb{R}, F(x) < 1\}.$$

We shall suppose that the expectation of X is finite ; i.e., $\mathbb{E}|X| < \infty$. Then, for any threshold u , the mean excess function $e(u)$ associated with X is defined (see, e.g., Kotz and Shanbhag, 1980, Hall and Wellner, 1981, Guess and Proschan, 1988) by

$$e(u) = \mathbb{E}(X - u | X > u) = \begin{cases} \frac{1}{\bar{F}(u)} \int_u^\infty \bar{F}(t) dt & \text{if } \bar{F}(u) > 0 \\ 0 & \text{whenever } \bar{F}(u) = 0, \end{cases} \quad (1)$$

where $\bar{F} = 1 - F$ represents the survival function of X . A natural way of estimating the mean excess function $e(u)$ is achieved by using the plug-in method ; that is, replacing the survival function in (1) by its empirical counterpart, as did Yang (1978).

Let X_1, \dots, X_n be a sequence of n independent copies of X . Then, the plug-in estimator of $e(u)$ is defined as

$$e_n(u) = \frac{\sum_{i=1}^n (X_i - u) \mathbb{I}_{[X_i > u]}}{\sum_{i=1}^n \mathbb{I}_{[X_i > u]}} = \frac{\sum_{i=1}^n X_i \mathbb{I}_{[X_i > u]}}{\sum_{i=1}^n \mathbb{I}_{[X_i > u]}} - u, \quad (2)$$

where $\mathbb{I}_{[X > u]} = 1$ if $X > u$ and 0 otherwise.

The asymptotic behavior of the estimator $e_n(u)$ has been investigated by some authors. Yang (1978) established the uniform strong consistency of $e_n(u)$ and the weak convergence of the process $Z_n(u) = \sqrt{n}[e_n(u) - e(u)]$ to a Gaussian process over fixed intervals of the form $[0, x_0]$, $x_0 < x_F$. Under some regularity conditions, Hall and Wellner (1981) extended these results to $[0, \infty)$. They also provided simultaneous confidence bands for the mean excess function $e(u)$ by transforming the limiting process to Brownian motion. Later on, Csörgő and Zitikis (1996) showed that these results does not hold, in general, on the half line $[0, \infty)$, and proposed to use weight functions to overcome the difficulties. Their approach also allows to obtain confidence bands for the mean excess function $e(u)$.

In this paper, we propose another approach based on the modern theory of empirical processes indexed by functions which enables us to study, with very mild conditions on the distribution function F , the asymptotic behavior of the estimator $e_n(u)$, uniformly in $u \in I = [u_0, u_1]$, with $u_0 < u_1 < x_F$.

The paper makes three contributions. First, by using the Vapnik-Chervonenkis (VC) classes and the entropy number techniques, we establish the uniform strong consistency of the empirical mean excess function $e_n(u)$ over any fixed interval I . Precisely, we show that for any u_1 less than x_F (the right-endpoint of the distribution F),

$$\sup_{u \leq u_1} |e_n(u) - e(u)| \rightarrow 0 \text{ a.s. as } n \rightarrow \infty.$$

Second, making use of the modern theory of weak convergence of empirical processes indexed by functions mainly exposed in (Van Der Vaart, 1996), we prove that, for any fixed interval I of the above type, the empirical mean excess process

$\sqrt{n}[e_n(u) - e(u)]$ indexed by a class \mathcal{F}_3 , weakly converges to a Gaussian process ; that is,

$$\{ \sqrt{n}[e_n(u) - e(u)], u \in I \} \xrightarrow{w} \{ \mathbb{G}(h_u), u \in I \},$$

where \mathbb{G} is a Gaussian centered functional process with a well defined covariance function and $\mathcal{F}_3 = \{h_u, u \in I\}$ is a class of functions to be precised later.

Third, combining Talagrand’s inequality (see Talagrand, 1994) and the Einmahl and Mason moment bound (see Einmahl, 2000), we arrive at our best achievement which consists of finding almost sure consistency bands for the mean excess function $e(u)$. Precisely, we establish that for any $\varepsilon > 0$, for all large n ,

$$\mathbb{P} \left(e_n(u) - \frac{E_n}{\sqrt{n}} < e(u) < e_n(u) + \frac{E_n}{\sqrt{n}}, u \in I \right) > 1 - \varepsilon,$$

where $(E_n)_{(n \geq 1)}$ is a non-random sequence of real numbers precised in Theorem 3.

To develop our methodology, we need some notations.

For all $u \leq x_F$, consider the real functions $f_u(x) = x\mathbb{I}_{[x>u]}$, $g_u(x) = \mathbb{I}_{[x>u]}$, $x \in \mathbb{R}$ and, define

$$\mathbb{P}_X(f_u) = \int_{\mathbb{R}} f_u(x)dF(x) = \int_{x>u} x dF(x)$$

and

$$\mathbb{P}_X(g_u) = \int_{\mathbb{R}} g_u(x)dF(x) = \int_{x>u} dF(x) = \bar{F}(u),$$

where \mathbb{P}_X denotes the probability law of X .

Let \mathbb{P}_n be the empirical measure associated with the sample X_1, \dots, X_n . Then, for all u , we have

$$\mathbb{P}_n(f_u) = \frac{1}{n} \sum_{i=1}^n X_i \mathbb{I}_{[X_i>u]} \text{ and } \mathbb{P}_n(g_u) = \frac{1}{n} \sum_{i=1}^n \mathbb{I}_{[X_i>u]},$$

and formulae (1) and (2) become

$$e(u) = \begin{cases} \frac{\mathbb{P}_X(f_u)}{\mathbb{P}_X(g_u)} - u & \text{if } u \leq x_F \\ 0 & \text{if } u > x_F. \end{cases}$$

and

$$e_n(u) = \begin{cases} \frac{\mathbb{P}_n(f_u)}{\mathbb{P}_n(g_u)} - u & \text{if } u \leq X_{n,n} \\ 0 & \text{if } u > X_{n,n}, \end{cases}$$

where $X_{n,n} = \max_{1 \leq i \leq n} X_i$.

Note that our consistency bands are asymptotically optimal in the sense that the coverage probability tends to 1, as $n \rightarrow \infty$ (see, Theorem 3). These results allow us to set graphical goodness-of-fitting test based on the empirical mean excess function, and to apply this test to Dow Jones data. We find that the Generalized hyperbolic family distribution reveals, itself, to be very adequate to generally fit financial data.

The remainder of the paper is organized as follows. We state uniform almost sure (*a.s*) convergence results in SECTION 2 and finite-distribution and functional normality theorems in SECTION 3. SECTION 4 is devoted to setting *a.s* consistency bands for the mean excess function $e(u)$. In SECTION 5, simulation studies and data driven applications using Dow Jones databases are provided. We finish the paper by a concluding section.

Before we go any further, it is worth mentioning that, in the sequel, all the suprema, taken over $u < u_1$, are measurable since the functions of u that we consider below, are left or right continuous. This means that we are in the pointwise-measurability scheme. Thus, even when we use the results and concepts in (Van Der Vaart, 1996), we do not need neither exterior and interior integrals, nor convergence in outer probability.

2. Almost Sure Convergence

In this section we are going to prove the uniform almost sure convergence of the empirical mean excess function by using Vapnik-Chervonenkis (VC) classes and bracketing numbers.

Theorem 1. *Suppose that $\mathbb{E}|g_u(X_1)| < \infty$ and $\mathbb{E}|f_u(X_1)| < \infty$, then*

$$\sup_{u < x_F} |\mathbb{P}_n(g_u) - \mathbb{P}_X(g_u)| \rightarrow 0 \text{ a.s. as } n \rightarrow \infty$$

and

$$\sup_{u < x_F} |\mathbb{P}_n(f_u) - \mathbb{P}_X(f_u)| \rightarrow 0 \text{ a.s. as } n \rightarrow \infty.$$

Moreover, if $\mathbb{E}X_1^2 < +\infty$, we have for any fixed $u_1 < x_F$,

$$\sup_{u \leq u_1} |e_n(u) - e(u)| \rightarrow 0 \text{ a.s. as } n \rightarrow \infty.$$

Proof. We notice that $\mathcal{F}_1 = \{g_u, u < x_F\}$ is a class of monotone real functions with values in $[0, 1]$. By Theorem 2.7.5 in (Van Der Vaart, 1996), the bracketing number $N_{[\cdot]}(\varepsilon, \mathcal{F}_1, L_r(Q))$ is finite (bounded by $\exp(K/\varepsilon)$, for every probability measure Q , any real $r \geq 1$, and a constant K that only depends on r). Since $\mathbb{E}|g_u(X_1)| < \infty$ for $u < x_F$, \mathcal{F}_1 is functional Glivenko-Cantelli class in the sense of Theorem 2.4.1 in (Van Der Vaart, 1996), meaning that

$$\sup_{u < x_F} |\mathbb{P}_n(g_u) - \mathbb{P}_X(g_u)| \rightarrow 0 \text{ a.s. as } n \rightarrow \infty. \tag{3}$$

The class $\mathcal{F}_2 = \{f_u, u \in [u_0, u_1]\}$, with $u_1 < x_F$, is a Vapnik-Chervonenkis class with index $V(\mathcal{F}_2) = 3$ and its envelop is $G = \max(|f_{u_0}(x)|, |f_{u_1}(x)|)$. Then it satisfies the uniform entropy condition 2.4.1 in (Van Der Vaart, 1996). Then \mathcal{F}_2 is a Donsker class and hence it is a Glivenko Cantelli class, that is

$$\sup_{u < x_F} |\mathbb{P}_n(f_u) - \mathbb{P}_X(f_u)| \rightarrow 0 \text{ a.s. as } n \rightarrow \infty. \tag{4}$$

To finish, fix $u_1 < x_F$. Then for $u \leq u_1$ and n large enough, we have

$$\begin{aligned} e_n(u) - e(u) &= \frac{\mathbb{P}_n(f_u)}{\mathbb{P}_n(g_u)} - \frac{\mathbb{P}_X(f_u)}{\mathbb{P}_X(g_u)} \\ &= (\mathbb{P}_n(g_u))^{-1}(\mathbb{P}_n(f_u) - \mathbb{P}_X(f_u)) - \mathbb{P}_X(f_u) \times \frac{\mathbb{P}_n(g_u) - \mathbb{P}_X(g_u)}{\mathbb{P}_n(g_u)\mathbb{P}_X(g_u)}. \end{aligned}$$

Then

$$|e_n(u) - e(u)| \leq |\mathbb{P}_n(g_u)|^{-1} \times |\mathbb{P}_n(f_u) - \mathbb{P}_X(f_u)| + |\mathbb{P}_X(f_u)| \times \frac{|\mathbb{P}_n(g_u) - \mathbb{P}_X(g_u)|}{|\mathbb{P}_n(g_u)\mathbb{P}_X(g_u)|}. \tag{5}$$

Let

$$\epsilon_n = \sup_{u < x_F} |\mathbb{P}_n(f_u) - \mathbb{P}_X(f_u)| \text{ and } \delta_n = \sup_{u < x_F} |\mathbb{P}_n(g_u) - \mathbb{P}_X(g_u)|. \tag{6}$$

From (3) and (4) above, we have

$$\epsilon_n \rightarrow 0 \text{ a.s. and } \delta_n \rightarrow 0 \text{ a.s., as } n \rightarrow \infty.$$

Now for $u \leq u_1$, we have $\mathbb{P}_X(g_u) \geq \mathbb{P}_X(g_{u_1})$ and from (6),

$$\begin{aligned} -\delta_n &\leq \mathbb{P}_n(g_u) - \mathbb{P}_X(g_u) \leq \delta_n \\ -\delta_n + \mathbb{P}_X(g_u) &\leq \mathbb{P}_n(g_u) \leq \delta_n + \mathbb{P}_X(g_u), \end{aligned}$$

since $\mathbb{P}_n(g_u) \geq \mathbb{P}_X(g_u) - \delta_n > 0$ for n large enough, then $(\mathbb{P}_n(g_u))^{-1} \leq (\mathbb{P}_X(g_u) - \delta_n)^{-1}$.

We also have

$$|\mathbb{P}_X(f_u)| = \left| \int_u x dF(x) \right| \leq \left| \int_{\mathbb{R}} x dF(x) \right| \leq \int_{\mathbb{R}} |x| dF(x) = \mathbb{E}|X| = \alpha < \infty. \tag{7}$$

Thus

$$\sup_{u \leq u_1} |e_n(u) - e(u)| \leq \epsilon_n [\mathbb{P}_X(g_{u_1}) - \delta_n]^{-1} + \alpha [\mathbb{P}_X(g_{u_1})(\mathbb{P}_X(g_{u_1}) - \delta_n)]^{-1} \delta_n$$

and then

$$\sup_{u \leq u_1} |e_n(u) - e(u)| \rightarrow 0 \text{ a.s. as } n \rightarrow \infty. \quad \square$$

3. Asymptotic normality of $e_n(u)$

In this section, we are concerned with weak laws of the empirical mean excess process $\{\sqrt{n}[e_n(u) - e(u)], u \in I\}$, where I is a given interval. Hereafter $\{\mathbb{G}(g), g \in \mathcal{G}\}$ denotes a Gaussian centered functional stochastic process with variance-covariance function

$$\Gamma(g_1, g_2) = \int (g_1(x) - \mathbb{E}g_1(X_1))(g_2(x) - \mathbb{E}g_2(X_1))dF(x).$$

Theorem 2. Let X_1, X_2, \dots be iid rv's with common finite second moments. Put $I = [u_0, u_1]$, with $u_0 < u_1 < x_F$ and define the functions of $t \in \mathbb{R}$,

$$h_u(t) = \mathbb{P}_X(g_u)^{-1} f_u(t) - \mathbb{P}_X(f_u) \mathbb{P}_X^{-2}(g_u) g_u(t), \text{ for } u \in I.$$

Suppose that F is continuous and satisfies

$$\limsup_{\delta \rightarrow 0} \sup_{(v, v-\delta) \in I^2} \left(\frac{F(v) - F(v-\delta)}{\sqrt{\delta}} \right)^2 = 0.$$

Then, the functional empirical processes $\{\mathbb{G}_n(g_u), u \in I\}$ and $\{\mathbb{G}_n(f_u), u \in I\}$ weakly converge respectively to $\{\mathbb{G}(g_u), u \in I\}$ and $\{\mathbb{G}(f_u), u \in I\}$ in $\ell^\infty(I)$. And consequently, the empirical mean excess process $\{\sqrt{n}[e_n(u) - e(u)], u \in I\}$ weakly converges to $\{\mathbb{G}(h_u), u \in I\}$, where $h_u = \frac{f_u}{\mathbb{P}_X(g_u)} - \frac{\mathbb{P}_X(f_u)}{\mathbb{P}_X^2(g_u)} g_u, u \in I$.

Before we give the proof of this theorem, we need the following lemma.

Lemma 1. Let g be a finite measurable function defined on \mathbb{R} such that $\mathbb{E}g(X_1)^2 < \infty$. Let $u_0 < u_1 < x_F$. Define for any fixed $v \in \mathbb{R}$ and $\delta > 0$

$$\sigma^2(v, \delta) = \int_{v-\delta}^v (g(x) - \mathbb{E}g(X))^2 dF(x).$$

Let for a fixed $n \geq 1, u \in \mathbb{R}$

$$S_n(u) = \frac{1}{\sqrt{n}} \sum_{j=1}^n [g(X_j) \mathbb{I}_{(X_j > u)} - \mathbb{E}g(X_j) \mathbb{I}_{(X_j > u)}].$$

If $\sup_{u_0 \leq v \leq u_1} \frac{\sigma^4(v, \delta)}{\delta} \rightarrow 0$ as $\delta \rightarrow 0$ and $\sup_{u_0 \leq x \leq u_1} |g(x) - \mathbb{E}g(X)| < \infty$,

then

$$\lim_{\delta \rightarrow 0} \sup_{u_0 \leq v \leq u_1} \sup_{n \geq 1} \frac{1}{\delta} P(\sup_{v-\delta \leq u \leq v} |S_n(u) - S_n(v)| \geq \eta) = 0.$$

Proof of Lemma 1. We fix $v \in \mathbb{R}$ and consider $\alpha = \sup_{v-\delta < u < v} |S_n(u) - S_n(v)|$. Observe that for $u < v$,

$$S_n(u) - S_n(v) = \frac{1}{\sqrt{n}} \sum_{j=1}^n \{g(X_j) \mathbb{I}_{[u, v]}(X_j) - \mathbb{E}g(X_j) \mathbb{I}_{[u, v]}(X_j)\}.$$

Given that for all $(u, v) \in \mathbb{R}^2$, we have

$$|S_n(v) - S_n(u)| \leq \frac{1}{\sqrt{n}} \sum_{j=1}^n [|g(X_j)| + |\mathbb{E}g(X_j)|] < \infty,$$

it comes that α is finite. So for any $\varepsilon > 0$, we can find $\bar{u} \in [v - \delta, v[$ such that,

$$|S_n(\bar{u}) - S_n(v)| \geq \alpha - \varepsilon. \tag{8}$$

Now, let $\delta > 0$. Define for any $p \geq 1$, and consider $u_j(p) = u_j = v - \delta + j\delta/p, j = 0, \dots, p$.

Let us prove that for $\varepsilon > 0$,

$$\lim_{p \rightarrow \infty} \max_{0 \leq j \leq p} |S_n(u_j) - S_n(v)| \geq \alpha - \varepsilon.$$

For each $p \geq 1$, let j such that

$$u_{j-1}(p) \leq \bar{u} \leq u_j(p).$$

We have,

$$\begin{aligned} |S_n(u_j) - S_n(v)| &\geq |S_n(\bar{u}) - S_n(v)| - \left| \frac{1}{\sqrt{n}} \sum_{i=1}^n (g(X_i)\mathbb{I}_{] \bar{u}, u_j(p)]}(X_i) - \mathbb{E}g(X_i)\mathbb{I}_{] \bar{u}, u_j(p)]}(X_i) \right| \\ &\geq |S_n(\bar{u}) - S_n(v)| - R_j(p), \end{aligned}$$

by denoting

$$R_j(p) = \left| \frac{1}{\sqrt{n}} \sum_{i=1}^n (g(X_i)\mathbb{I}_{] \bar{u}, u_j(p)]}(X_i) - \mathbb{E}g(X_i)\mathbb{I}_{] \bar{u}, u_j(p)]}(X_i) \right|.$$

We get from (8)

$$\max_{0 \leq j \leq p} |S_n(u_j) - S_n(v)| \geq \alpha - \varepsilon - R_j(p).$$

For a fixed $n \geq 1, R_j(p) \rightarrow 0$ as $p \rightarrow \infty$, since the sequence of intervals $(] \bar{u}, u_j(p)])$ _{$p \geq 1$} decreases to the empty set as $p \rightarrow \infty$.

Next, consider the collection of the points $\{u_j(\ell), 0 \leq j \leq p, 1 \leq \ell \leq p\}$ and denote the set of its distinct values between them as $\{\bar{u}_j, 1 \leq j \leq m(p)\}$. We still have $|\bar{u}_j - \bar{u}_{j-1}| \leq \delta/p$, and we surely have for any $\varepsilon > 0$

$$\lim_{p \rightarrow \infty} \max_{0 \leq j \leq m(p)} |S_n(\bar{u}_j) - S_n(v)| \geq \alpha - \varepsilon$$

and then

$$\lim_{p \rightarrow \infty} \max_{0 \leq j \leq m(p)} |S_n(\bar{u}_j) - S_n(v)| \geq \alpha$$

and finally

$$\sup_{p \geq 1} \max_{0 \leq j \leq m(p)} |S_n(\bar{u}_j) - S_n(v)| = \alpha.$$

By construction, $\max_{0 \leq j \leq m(p)} |S_n(\bar{u}_j) - S_n(v)|$ is non decreasing in p . So, by the Monotone Convergence Theorem, for any fixed $v > 0$ and for any $\eta > 0$,

$$\mathbb{P} \left(\sup_{v-\delta \leq u \leq v} |S_n(u) - S_n(v)| \geq \eta \right) = \lim_{p \uparrow \infty} \mathbb{P} \left(\max_{1 \leq j \leq m(p)} |S_n(\bar{u}_j) - S_n(v)| \geq \eta \right). \tag{9}$$

Put $Z_h = \sum_{i=1}^n (g(X_i)\mathbb{I}_{] \bar{u}_{h-1}, \bar{u}_h]}(X_i) - \mathbb{E}g(X_i)\mathbb{I}_{] \bar{u}_{h-1}, \bar{u}_h]}(X_i)), h \geq 1$.

We have

$$\sqrt{n}(S_n(\bar{u}_j) - S_n(v)) = \sum_{h=j}^{m(p)} Z_h = T_{m(p)-j}$$

with

$$\sqrt{n}(S_n(v - \delta) - S_n(v)) = \sum_{i=1}^{m(p)} Z_i = T_{m(p)} = T(n, u, \delta).$$

We observe that $\{T_1, T_2, \dots, T_{m(p)}\}$ are partial sums of i.i.d. centered random variables so that the T_j^4 form a submartingale. By the maximal inequality form submartingales, for any fixed p

$$\begin{aligned} \mathbb{P}\left(\max_{1 \leq j \leq m(p)} |S_n(\bar{u}_j) - S_n(v)| \geq \eta\right) &= \mathbb{P}\left(\max_{1 \leq j \leq m(p)} |T_j| \geq \eta \sqrt{n}\right) \leq \frac{1}{\eta^4 n^2} \mathbb{E}T_{m(p)}^4 \\ &\leq \frac{1}{\eta^4 n^2} \mathbb{E}T(n, u, \delta)^4. \end{aligned}$$

Since the right hand side does not depend on p , we get by (9)

$$\frac{1}{\delta} \mathbb{P}\left(\sup_{v-\delta \leq u \leq v} |S_n(u) - S_n(v)| \geq \eta\right) \leq \frac{1}{\delta \eta^4 n^2} \mathbb{E}T(n, u, \delta)^4.$$

Notice that $T(n, u, \delta)$ is a sum of n i.i.d centered random variables with variance

$$\kappa_1(v, \delta) = \sigma^2(v, \delta) = \int_{v-\delta}^v (g(x) - \mathbb{E}(g(x)))^2 dF(x)$$

and fourth moment

$$\kappa_2(v, \delta) = \int_{v-\delta}^v (g(x) - \mathbb{E}(g(x)))^4 dF(x).$$

Simple computations give (see the APPENDIX 7.1 for a simple proof of that)

$$\mathbb{E}(T(n, u, \delta))^4 = n\kappa_2(v, \delta) + 3n(n-1)\kappa_1^2(v, \delta).$$

By putting these facts together, we arrive at

$$\begin{aligned} \frac{1}{\delta} \mathbb{P}\left(\sup_{v-\delta \leq u \leq v} |S_n(u) - S_n(v)| \geq \eta\right) &\leq \eta^{-4} \left(\frac{n\kappa_2(v, \delta) + 3n(n-1)\sigma^4(v, \delta)}{\delta n^2}\right) \\ &\leq \eta^{-4} \left(\frac{\kappa_2(v, \delta)}{\delta} \times \frac{1}{n} + 3 \frac{\sigma^4(v, \delta)}{\delta} \times \left[1 - \frac{1}{n}\right]\right). \end{aligned}$$

Remark that

$$\begin{aligned} \sup_{u_0 \leq v \leq u_1} \frac{\kappa_2(v, \delta)}{\delta} &\leq \left(\sup_{u_0 \leq x \leq u_1} |g(x) - \mathbb{E}g(X)|\right)^4 \times \delta^{-1} \times \sup_{u_0 \leq v \leq u_1} \int_{v-\delta}^v dF(x) \\ &\leq \left(\sup_{u_0 \leq x \leq u_1} |g(x) - \mathbb{E}g(X)|\right)^4 \times \sup_{u_0 \leq v \leq u_1} \frac{F(v) - F(v-\delta)}{\delta}. \end{aligned}$$

We finally get

$$\lim_{\delta \rightarrow 0} \sup_{u_0 \leq v \leq u_1} \sup_{n \geq 1} \frac{1}{\delta} \mathbb{P}\left(\sup_{v-\delta \leq u \leq v} |S_n(u) - S_n(v)| \geq \eta\right) = 0$$

whenever $\lim_{\delta \rightarrow 0} \sup_{u_0 \leq v \leq u_1} \frac{\sigma^4(v, \delta)}{\delta} = 0$ and $\sup_{u_0 \leq x \leq u_1} |g(x) - \mathbb{E}g(X)| < +\infty$.

This achieves the proof of the lemma. \square

Proof of Theorem 2.

By Theorem 2.7.5 in (Van Der Vaart, 1996) applied to \mathcal{F}_1 and by the fact that \mathcal{F}_2 is a Vapnik-Chervonenkis class, condition (2.5.1) is satisfied for both \mathcal{F}_1 and \mathcal{F}_2 thus \mathcal{F}_1 and \mathcal{F}_2 are Donsker classes. This may be used in a simple manner to get

$$A_n = \max(\sup_{u \in I} |\mathbb{G}_n(g_u)|, \sup_{u \in I} |\mathbb{G}_n(f_u)|) = O_{\mathbb{P}}(1, I) \text{ as } n \rightarrow \infty. \tag{10}$$

Denote the functional empirical process for any real function g by

$$\mathbb{G}_n(g) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \{g(X_i) - \mathbb{E}g(X_i)\}.$$

Remind that for any Donsker class \mathcal{G} , the functional stochastic process $\{\mathbb{G}_n(g), g \in \mathcal{G}\}$ converges in law to a Gaussian and centered stochastic process $\{\mathbb{G}(g), g \in \mathcal{G}\}$ whose variance-covariance function is

$$\Gamma(g_1, g_2) = \int (g_1(x) - \mathbb{E}g(X_1))(g_2(x) - \mathbb{E}g_2(X_1))dF(x).$$

We have, as $n \rightarrow \infty$

$$\begin{aligned} \mathbb{P}_n(g_u) &= \mathbb{P}_X(g_u) + \frac{\mathbb{G}_n(g_u)}{\sqrt{n}} \\ \mathbb{P}_n(f_u) &= \mathbb{P}_X(f_u) + \frac{\mathbb{G}_n(f_u)}{\sqrt{n}}. \end{aligned}$$

Thus

$$\begin{aligned} \sqrt{n}(e_n(u) - e(u)) &= \sqrt{n}\left(\frac{\mathbb{P}_n(f_u)}{\mathbb{P}_n(g_u)} - \frac{\mathbb{P}_X(f_u)}{\mathbb{P}_X(g_u)}\right) \\ &= \frac{1}{\mathbb{P}_n(g_u)} \sqrt{n}(\mathbb{P}_n(f_u) - \mathbb{P}_X(f_u)) - \mathbb{P}_X(f_u) \frac{\sqrt{n}(\mathbb{P}_n(g_u) - \mathbb{P}_X(g_u))}{\mathbb{P}_n(g_u)\mathbb{P}_X(g_u)} \\ &= \frac{1}{\mathbb{P}_n(g_u)} \left[\mathbb{G}_n(f_u) - \frac{\mathbb{P}_X(f_u)}{\mathbb{P}_X(g_u)} \mathbb{G}_n(g_u) \right] \\ &= \frac{1}{\mathbb{P}_n(g_u)} \left[\mathbb{G}_n\left(f_u - \frac{\mathbb{P}_X(f_u)}{\mathbb{P}_X(g_u)} g_u\right) \right]. \end{aligned}$$

We find

$$\begin{aligned} (\mathbb{P}_n(g_u))^{-1} &= \left[\mathbb{P}_X(g_u) + \frac{\mathbb{G}_n(g_u)}{\sqrt{n}} \right]^{-1} \\ &= \mathbb{P}_X^{-1}(g_u) \left[1 + \mathbb{P}_X^{-1}(g_u) \times n^{-1/2} \times \mathbb{G}_n(g_u) \right]^{-1} \\ &= \mathbb{P}_X^{-1}(g_u) \left[1 - \mathbb{P}_X^{-1}(g_u) \times n^{-1/2} \times \mathbb{G}_n(g_u) + \mathbb{P}_X^{-1}(g_u) \times \theta(n^{-1/2} \times \mathbb{G}_n(g_u)) \right] \end{aligned}$$

Since \mathcal{F}_1 is a Donsker class, then $\sup_{u \in I} |\mathbb{G}_n(g_u)| = \|\mathbb{G}_n\|_{\mathcal{F}_1} = O_{\mathbb{P}}(1, I)$. So

$$(\mathbb{P}_n(g_u))^{-1} = \mathbb{P}_X^{-1}(g_u) \left[1 - \mathbb{P}_X^{-1}(g_u) \times n^{-1/2} \times O_{\mathbb{P}}(1, I) \right]$$

Let us remind that $h_u = \mathbb{P}_X(g_u)^{-1} f_u - \mathbb{P}_X(f_u) \mathbb{P}_X^{-2}(g_u) g_u$. Then, for $u \in I$, we get

$$\begin{aligned} \sqrt{n}(e_n(u) - e(u)) &= \left[\mathbb{G}_n\left(f_u - \frac{\mathbb{P}_X(f_u)}{\mathbb{P}_X(g_u)} g_u\right) \right] \times \left[\mathbb{P}_X^{-1}(g_u) - \mathbb{P}_X^{-2}(g_u) \times n^{-1/2} \times O_{\mathbb{P}}(1, I) \right] \\ &= \mathbb{G}_n(h_u) + \mathbb{G}_n(h_u) \times \mathbb{P}_X^{-1}(g_u) \times n^{-1/2} \times O_{\mathbb{P}}(1, I). \end{aligned}$$

We finally have

$$\sqrt{n}(e_n(u) - e(u)) = \mathbb{G}_n(h_u) + \mathbb{G}_n(h_u) \times o_{\mathbb{P}}(1, I). \quad \square \tag{11}$$

Lemma 2. The class $\mathcal{F}_3 = \left\{ h_u = \frac{f_u}{\mathbb{P}_X(g_u)} - \frac{\mathbb{P}_X(f_u)}{\mathbb{P}_X^2(g_u)} g_u, u \in I \right\}$ is a Donsker Class.

Proof. At this step, we want to prove that $\mathcal{F}_3 = \{h_u, u_0 \leq u \leq u_1\}$ is a Donsker Class. Since we obviously have, by the Central Limit Theorem, finite distribution convergence of $\{\mathbb{G}_n(h_u), u \in I\}$ to the stochastic process $\{\mathbb{G}(h_u), u \in I\}$ in $\ell^\infty(\mathcal{F}_3)$, we only need to prove the asymptotic tightness of $\{\mathbb{G}_n(h_u), u \in I\}$.

In view of Theorem in 1.5.7 in (Van Der Vaart, 1996), it is enough to prove that

$$\lim_{\delta \rightarrow 0} \sup_{u \in I} \limsup_{n \rightarrow \infty} \frac{1}{\delta} \mathbb{P}\left(\sup_{v - \delta \leq u \leq v} |\mathbb{G}_n(h_u) - \mathbb{G}_n(h_v)| \geq \eta \right) = 0.$$

Here, we apply **Lemma 1** for the nondecreasing measurable function $g(x) = x$ and $g(x) = 1$. In both cases, we inspect the assumptions of this lemma and see that if $g(x) = x$, we get $g(x) \leq g(u_1) = u_1$ for any $u_0 \leq x \leq u_1$ and thus

$$\begin{aligned} \sup_{u_0 \leq v \leq u_1} \frac{\sigma^4(v, \delta)}{\delta} &= \sup_{u_0 \leq v \leq u_1} \frac{1}{\delta} \left(\int_{v-\delta}^v (g(x) - \mathbb{E}g(x))^2 dF(x) \right)^2 \\ &\leq |u_1 - \mathbb{E}(X)|^4 \times \sup_{u_0 \leq v \leq u_1} \left(\frac{F(v) - F(v - \delta)}{\sqrt{\delta}} \right)^2 \rightarrow 0 \text{ as } \delta \rightarrow 0, \end{aligned}$$

and

$$\sup_{x \in I} |g(x) - \mathbb{E}g(X)| \leq |u_1| + |\mathbb{E}(X)| < \infty.$$

If $g(x) = 1$, the result is obvious.

We can apply **Lemma 1** and we will get,

$$\limsup_{\delta \rightarrow 0} \limsup_{u \in I} \limsup_{n \rightarrow \infty} \frac{1}{\delta} \mathbb{P} \left(\sup_{v-\delta \leq u \leq v} |\mathbb{G}_n(f_u) - \mathbb{G}_n(f_v)| \geq \eta \right) = 0$$

and

$$\limsup_{\delta \rightarrow 0} \limsup_{u \in I} \limsup_{n \rightarrow \infty} \frac{1}{\delta} \mathbb{P} \left(\sup_{v-\delta \leq u \leq v} |\mathbb{G}_n(g_u) - \mathbb{G}_n(g_v)| \geq \eta \right) = 0.$$

But by Theorem 8.3 of Billingsley (1968), p.56, and by Theorem 2.2 in Lo (2014), these two previous equalities entail, that

$$\limsup_{\delta \rightarrow 0} \limsup_{u \in I} \limsup_{n \rightarrow \infty} \mathbb{P} \left(\sup_{|u-v| \leq \delta, (u,v) \in I^2} |\mathbb{G}_n(f_u) - \mathbb{G}_n(f_v)| \geq \eta \right) = 0$$

and

$$\limsup_{\delta \rightarrow 0} \limsup_{u \in I} \limsup_{n \rightarrow \infty} \mathbb{P} \left(\sup_{|u-v| \leq \delta, (u,v) \in I^2} |\mathbb{G}_n(g_u) - \mathbb{G}_n(g_v)| \geq \eta \right) = 0.$$

Next, we use the following development for $(u, v) \in I^2$

$$\begin{aligned} h_u - h_v &= \mathbb{P}_X^{-1}(g_u) \left(f_u - \frac{\mathbb{P}_X(f_u)}{\mathbb{P}_X(g_u)} g_u \right) - \mathbb{P}_X^{-1}(g_v) \left(f_v - \frac{\mathbb{P}_X(f_v)}{\mathbb{P}_X(g_v)} g_v \right) \\ &= \underbrace{\mathbb{P}_X^{-1}(g_u) f_u - \mathbb{P}_X^{-1}(g_v) f_v}_{a(u, v)} - \underbrace{\left(\frac{\mathbb{P}_X(f_u)}{\mathbb{P}_X^2(g_u)} g_u - \frac{\mathbb{P}_X(f_v)}{\mathbb{P}_X^2(g_v)} g_v \right)}_{b(u, v)} \\ &= a(u, v) - b(u, v). \end{aligned}$$

We get

$$\begin{aligned} a(u, v) &= \mathbb{P}_X^{-1}(g_u) f_u - \mathbb{P}_X^{-1}(g_u) f_v + \mathbb{P}_X^{-1}(g_u) f_v - \mathbb{P}_X^{-1}(g_v) f_v \\ &= \frac{f_u - f_v}{\mathbb{P}_X(g_u)} - \frac{\mathbb{P}_X(g_u) - \mathbb{P}_X(g_v)}{\mathbb{P}_X(g_u) \times \mathbb{P}_X(g_v)} \times f_v. \end{aligned}$$

Then

$$|\mathbb{G}_n(a(u, v))| \leq \frac{1}{\mathbb{P}_X(g_u)} \times |\mathbb{G}_n(f_u - f_v)| + \frac{|\mathbb{P}_X(g_u) - \mathbb{P}_X(g_v)|}{\mathbb{P}_X(g_u) \times \mathbb{P}_X(g_v)} \times \mathbb{G}_n(f_v).$$

Next

$$\begin{aligned} b(u, v) &= \frac{\mathbb{P}_X(f_u)}{\mathbb{P}_X^2(g_u)} \times g_u - \frac{\mathbb{P}_X(f_u)}{\mathbb{P}_X^2(g_u)} \times g_v + \frac{\mathbb{P}_X(f_u)}{\mathbb{P}_X^2(g_u)} \times g_v - \frac{\mathbb{P}_X(f_v)}{\mathbb{P}_X^2(g_v)} \times g_v \\ &= (g_u - g_v) \times \frac{\mathbb{P}_X(f_u)}{\mathbb{P}_X^2(g_u)} + \left[\frac{\mathbb{P}_X(f_u)}{\mathbb{P}_X^2(g_u)} - \frac{\mathbb{P}_X(f_v)}{\mathbb{P}_X^2(g_v)} \right] \times g_v. \end{aligned}$$

Next,

$$\begin{aligned} \frac{\mathbb{P}_X(f_u)}{\mathbb{P}_X^2(g_u)} - \frac{\mathbb{P}_X(f_v)}{\mathbb{P}_X^2(g_v)} &= \frac{\mathbb{P}_X(f_u)}{\mathbb{P}_X^2(g_u)} - \frac{\mathbb{P}_X(f_u)}{\mathbb{P}_X^2(g_v)} + \frac{\mathbb{P}_X(f_u)}{\mathbb{P}_X^2(g_v)} - \frac{\mathbb{P}_X(f_v)}{\mathbb{P}_X^2(g_v)} \\ &= \left(\frac{(\mathbb{P}_X(g_v) + \mathbb{P}_X(g_u)) \times (\mathbb{P}_X(g_v) - \mathbb{P}_X(g_u))}{\mathbb{P}_X^2(g_u) \times \mathbb{P}_X^2(g_v)} \right) \times \mathbb{P}_X(f_u) \\ &\quad + (\mathbb{P}_X(f_u) - \mathbb{P}_X(f_v)) \times \frac{1}{\mathbb{P}_X^2(g_v)}. \end{aligned}$$

Also,

$$\begin{aligned}
 |\mathbb{G}_n(b(u, v))| \leq & \frac{|\mathbb{P}_X(f_u)|}{\mathbb{P}_X^2(g_u)} \times |\mathbb{G}_n(g_u - g_v)| \\
 & + \frac{|\mathbb{P}_X(g_v) + \mathbb{P}_X(g_u)| \times |\mathbb{P}_X(g_v) - \mathbb{P}_X(g_u)|}{|\mathbb{P}_X^2(g_u) \times \mathbb{P}_X^2(g_v)|} \times |\mathbb{P}_X(f_u)| \times |\mathbb{G}_n(g_v)| \\
 & + |\mathbb{P}_X(f_u) - \mathbb{P}_X(f_v)| \times |\mathbb{G}_n(g_v)| \times \frac{1}{\mathbb{P}_X^2(g_v)}.
 \end{aligned}$$

For $(u, v) \in I^2$, let us use the bounds for $\mathbb{P}_X^{-1}(g_u)$, $\mathbb{P}_X^{-1}(g_v)$, and for $\mathbb{P}_X(f_u)$.

Precisely, we have $\mathbb{P}_X^{-1}(g_u) \leq (\bar{F}(u_1))^{-1}$, $\mathbb{P}_X^{-1}(g_v) \leq (\bar{F}(u_1))^{-1}$, and finally, from (7), we get $|\mathbb{P}_X(f_u)| \leq \mathbb{E}|X|$.

Thus, by using these bounds and (10), it comes that

$$\begin{aligned}
 \sup_{|u-v| \leq \delta, (u,v) \in I^2} |\mathbb{G}_n(h_u - h_v)| \leq & B_1 \times \sup_{|u-v| \leq \delta, (u,v) \in I^2} |\mathbb{G}_n(f_u - f_v)| + B_2 \times \sup_{|u-v| \leq \delta, (u,v) \in I^2} |\mathbb{G}_n(g_u - g_v)| \\
 & + \left(B_3 \times \sup_{|u-v| \leq \delta, (u,v) \in I^2} |\mathbb{P}_X(g_u) - \mathbb{P}_X(g_v)| + B_4 \times \sup_{|u-v| \leq \delta, (u,v) \in I^2} |\mathbb{P}_X(f_u) - \mathbb{P}_X(f_v)| \right) A_n,
 \end{aligned}$$

where

$$\begin{cases}
 B_1 = (\bar{F}(u_1))^{-1}; \\
 B_2 = \mathbb{E}|X| \times (\bar{F}(u_1))^{-2}; \\
 B_3 = \bar{F}(u_1)^{-2} \left((\bar{F}(u_1))^{-2} + 2\bar{F}(u_0) \right); \\
 B_4 = (\bar{F}(u_1))^{-2}; \\
 A_n = \max(\sup_{u \in I} |\mathbb{G}_n(g_u)|, \sup_{u \in I} |\mathbb{G}_n(f_u)|).
 \end{cases}$$

Now we observe that

$$\sup_{|u-v| \leq \delta, (u,v) \in I^2} |\mathbb{P}_X(g_u) - \mathbb{P}_X(g_v)| = \sup_{|u-v| \leq \delta, (u,v) \in I^2} |F(u) - F(v)|$$

and

$$\begin{aligned}
 \sup_{|u-v| \leq \delta, (u,v) \in I^2} |\mathbb{P}_X(f_u) - \mathbb{P}_X(f_v)| & \leq \sup_{|u-v| \leq \delta, (u,v) \in I^2} \left| \int_u^\infty t dF(t) - \int_v^\infty t dF(t) \right| \\
 & \leq \max(|u_0|, |u_1|) \sup_{|u-v| \leq \delta, (u,v) \in I^2} |F(u) - F(v)|.
 \end{aligned}$$

These quantities go to zero whenever F is continuous and hence uniformly continuous in I . Putting all these facts together and using (10) yield

$$\sup_{n \geq 1} \sup_{|u-v| \leq \delta, (u,v) \in I^2} |\mathbb{G}_n(h_u - h_v)| \rightarrow 0 \text{ as } \delta \rightarrow 0.$$

Finally \mathcal{F}_3 is a Donsker class, thus $\sup_{u \in I} |\mathbb{G}_n(h_u)| = O_{\mathbb{P}}(1, I)$ and we get from (11) that

$$\sqrt{n}(e_n(u) - e(u)) = \mathbb{G}_n(h_u) + o_{\mathbb{P}}(1, I).$$

This completes the proof. \square

Now we are going to concentrate on consistency bands for the mean excess function.

4. Consistency Bands

In this section we use uniform bounds of functional empirical processes based on Talagrand’s inequality (see Talagrand, 1994) and the Einmahl and Mason (2000) moment bound to construct almost surely consistency bands for the mean excess function $e(u)$ as follows.

Theorem 3. *Let X_1, X_2, \dots , be i.i.d random variables with finite second moments. Put $I = [u_0, u_1]$, with $-\infty < u_0 < u_1 < x_F$ and suppose that F is continuous and satisfies*

$$\limsup_{\delta \rightarrow 0} \sup_{(v, v-\delta) \in I^2} \left(\frac{F(v) - F(v - \delta)}{\sqrt{\delta}} \right)^2 = 0. \tag{12}$$

Then for any $\varepsilon > 0$, there exists n_0 such that for $n \geq n_0$,

$$\mathbb{P}\left(e_n(u) - \frac{E_n}{\sqrt{n}} < e(u) < e_n(u) + \frac{E_n}{\sqrt{n}}, u \in I\right) \geq 1 - \varepsilon,$$

with

$$E_n = \frac{1}{\overline{F}(u_1) - D_1/\sqrt{n}} \left(D_2 + \frac{D_1 \times \mathbb{E}|X|}{\overline{F}(u_1)} \right), \tag{13}$$

and where

$$\begin{cases} D_1 = 2AA_1 \sqrt{\log 2} + A_1 \\ D_2 = A A_1 M_1 \sqrt{\log M_1} + A_1, \\ M_1 = \max(2, \max(|u_0|, |u_1|)) \end{cases}$$

A and A_1 are universal constants.

The proof of this theorem is rather technical, so we postpone it to the APPENDIX SUBSECTION 7.2, where we also state the fundamental Talagrand’s inequality.

Remark : Condition (12) is quite very mild and is satisfied by most of the continuous usual distribution functions. Indeed, if F is absolutely continuous with respect to the Lebesgue measure, then F admits a density, say, f . By applying the *mean value theorem*, we get

$$\left(\frac{F(v) - F(v - \delta)}{\sqrt{\delta}} \right)^2 \leq \delta \times \sup_{x \in [v-\delta, v]} f^2(x).$$

But $\sup_{x \in [v-\delta, v]} f^2(x) < \infty$ because of the local continuity of the density f ; that is condition (12) holds. Thus, our proposed consistency bands can be constructed for a large number of continuous distribution functions. All the examples in SECTION 5, that are devoted to simulation studies, satisfy (12) through this argument.

Now, we are going to focus on the applications of our results.

5. Simulations and Applications

In this section, we are coming back on the empirical mean excess function (*emef* for short) to construct graphical tools for goodness-of-fit test.

In SUBSECTION 5.1, we present a simulation study. For this we consider some parametric distributions, for which we plot the *emef* and the consistency bands. For each distribution model, we generate $B = 6000$ samples of size $n = 4000$, and we compute the average mean excess function of these B samples. The graphs of these averages of these empirical mean excess functions would serve as stallions in the following sense : each other sample having an alike *emef* will suggest such an underlying distribution. We will also simulate our results with a special guest distribution : the generalized hyperbolic (*Gh* for short) family of distributions. Nowadays, this family is very important in financial modeling.

In SUBSECTION 5.2, we will use the *emef* graphs as stallions for real data. We shall focus on monthly returns and log-returns of Dow Jones data (03/01/2000 to 02/12/2011), and will find that these data strongly suggest *Gh* model.

This section, beyond financial data, shows how to use the *emef* for goodness-of-fit testing purposes. It opens a great variety of applications for different types of data.

5.1 Simulations

To assess the performance of our consistency bands, we present a simulation study. We draw simulated *emefs* for standard distributions and next for *Gh* family of distribution functions

5.1.1 *Emef* for Standard Distributions

We consider some simple distribution models listed in the Table 1 below, where the parameters are specified and the *emef* figures corresponding to each model are displayed.

Table 1. The *emef* for standard distributions with their specified parameters

Distributions	Parameters	Figures
<i>GPD</i>	$\xi = 0.25, \beta = 1$	Figure 1
	$\xi = -0.75, \beta = 1$	
Pareto	$\alpha = 7, \lambda = 3$	Figure 2
Exponential	$\lambda = 2$	
Gamma	$\alpha = 2, \beta = 0.001$	Figure 3
Beta	$\lambda = 7, \beta = 2$	

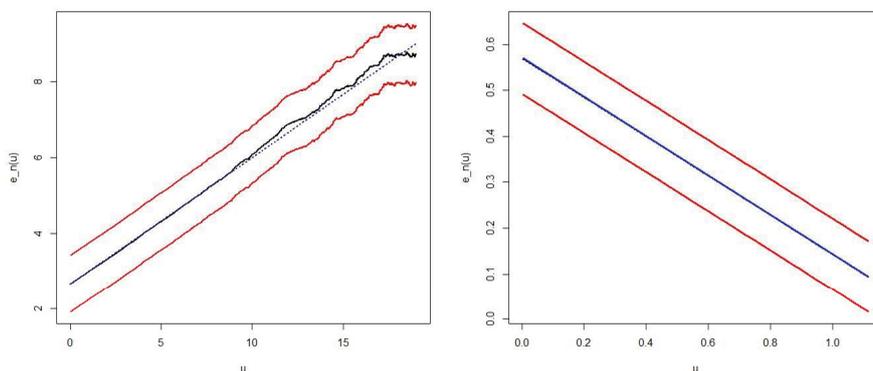


Figure 1. The *emef* for a generalized Pareto distributions $\xi = 0.25, \beta = 1$ and the right panel concerns the one with the parameters $\xi = -0.75, \beta = 1$.

5.1.2 Generalized Hyperbolic Models

Here, we consider the *emefs* of the *Gh* models. We need some definitions. The Lebesgue density function of the one dimensional *Gh* is given by

$$f_{\lambda, \alpha, \beta, \delta, \mu}(x) = \mathbf{a}_{(\lambda, \alpha, \beta, \delta, \mu)} \times (\delta^2 + (x - \mu)^2)^{(\lambda - \frac{1}{2})/2} \times e^{\beta(x - \mu)} \times K_{\lambda - \frac{1}{2}}(\alpha \sqrt{\delta^2 + (x - \mu)^2})$$

where $\mathbf{a}_{(\lambda, \alpha, \beta, \delta, \mu)} = \frac{(\alpha^2 - \beta^2)^{\frac{\lambda}{2}}}{\sqrt{2\pi} \alpha^{(\lambda - \frac{1}{2})} \delta^\lambda K_\lambda(\delta \sqrt{\alpha^2 - \beta^2})}$ is a norming constant to make the curve area equal to 1 and

$$K_\lambda(x) = \frac{1}{2} \int_0^\infty y^{\lambda - 1} \exp\left(-\frac{1}{2}x(y + y^{-1})\right) dy, \quad (x > 0)$$

is the modified Bessel function of the third kind with index λ .

The dependence of the parameters $\lambda, \alpha, \beta, \delta,$ and μ is as follows : $\alpha > 0$ determines the *shape*, $0 \leq |\beta| < \alpha$ the *skewness*, $\mu \in \mathbb{R}$ is a *location* parameter and $\delta > 0$ serves for *scaling*. The parameter $\lambda \in \mathbb{R}$ specifies the order K_λ function Bessel that appears in the *Gh* density function and may be used to obtain different subclasses of *Gh* distribution. We summarize the different possible domains for the parameters :

- If $\lambda < 0$ then $(\delta > 0, |\beta| \leq \alpha)$,
- If $\lambda = 0$ then $(\delta > 0, |\beta| < \alpha)$
- If $\lambda > 0$, then $(\delta \geq 0, |\beta| < \alpha)$.

An important characteristic of the *Gh* family is that, it embraces many special cases such as : Hyperbolic ($\lambda = 1$), Student-*t* ($\lambda < 0$), Variance Gamma ($\lambda > 0$), and the Normal Inverse Gaussian (NIG) ($\lambda = -0.5$) distributions. It nests

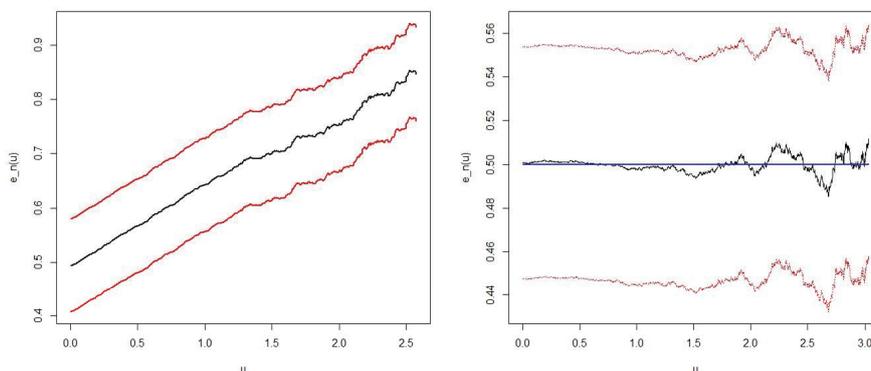


Figure 2. The left panel is the *emef* for a Pareto distribution with the parameters $\alpha = 7$ and $\lambda = 3$ and The *emef* for an Exponential distribution with the parameter $\lambda = 2$.

the Generalized Inverse Gamma (GIG) distribution defined only by the three parameters λ, α , and β . An Inverse Gaussian (IG) distribution is a GIG distribution with $\lambda = -0.5$ and a Gamma (Γ) distribution is also a GIG distribution with $\beta = 0$.

The *Gh* family contains also some limiting distributions such as Cauchy distributions with parameters μ and δ (obtained for $\lambda = -0.5$ and $\alpha = \beta = 0$). The Gaussian distribution with mean μ and variance σ^2 is obtained for $\lambda = -0.5$, for $\alpha, \delta \rightarrow \infty$ and $\frac{\delta}{\alpha} \rightarrow \sigma^2$. The Skew-Student *t* with ν degrees of freedom is obtained if $\alpha = |\beta|$, then $\nu = -2\lambda > 0$. The Student *t* distribution is obtained for $\alpha = \beta = 0, \mu = 0$ and $\delta = \sqrt{\nu}$. In the special case of hyperbolic distributions ($\lambda = 1$), we obtain the skewed Laplace distribution for $\delta = 0$.

All of these distributions have been used to model financial returns and log-returns data. In Table 2, we consider two specific *Gh* distributions with specified parameters. The corresponding *emef* plots are displayed in the column Figures.

Table 2. Specific and limiting *Gh* distributions with their specified parameters.

Distributions	Specified parameters					Figures
	λ	α	β	δ	μ	
t-stud. ^{spe}	-1.278	0.01186	0.01186	0.0766	1.005	Figure 4
	-1.247	0.0148	-0.0147	0.076	1.005	
NIG	-0.5	8.03	-1.37	0.051	0.0105	Figure 5
	-0.5	7.6	-1.24	0.052	0.0103	

5.2 Applications

We are now in a position to use the *emef* graphs already drawn to fit graphically the Dow Jones data. *Emefs* for Normal Inverse Gaussian (NIG) and *t*-student distributions are not monotonic functions. They decrease and increase like for the *emef* of returns and log-returns data. For this reason, we fit them to both monthly returns and log-returns from Dow Jones data base (see FIGURE 7, FIGURE 9, FIGURE 11, and FIGURE 13).

Dow Jones data base consists of several companies like AXP(American Express company), CSCO(Cisco Systems), DAX, CAT, IBM and so one. Each one having 5 values : from opening (*op*) values to closing (*cl*) values , also minimum (*min*), maximum (*max*), and volume (*vol*) values.

We select AXP and CSCO companies and we consider returns and log-returns for their values as shown in the Table 3. Then we construct their *emef* plots and their fitted counterparts. Estimated parameters and the *emef* plots are given in Table 4.

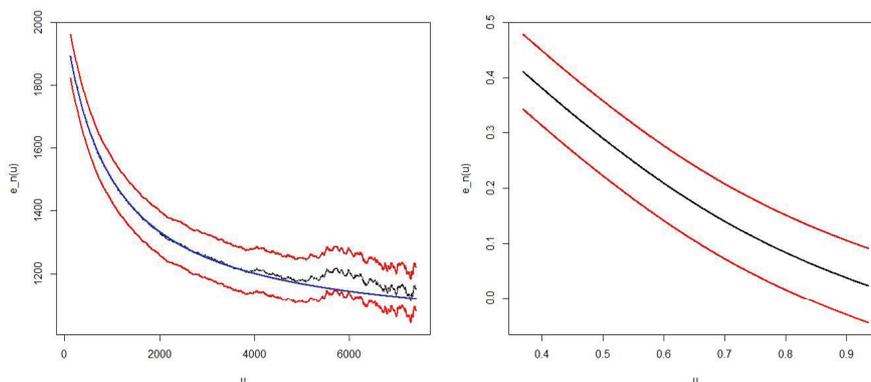


Figure 3. The *emef* for the Gamma distribution with the parameters $\alpha = 2, \beta = 0.001$ and the right one is the *emef* for the Beta distribution with the parameters $\lambda = 7, \beta = 2$.

Table 3. Real and Fitted *Emef* for AXP and CSCO companies.

Companies	Nature	Values	<i>Emef</i> plots	
			<i>Real emef</i>	<i>Fitted emef</i>
AXP	Returns	$\frac{op}{min}$	Figure 6	Figure 7
	Log-ret.	$\frac{max}{cl}$	Figure 8	Figure 9
CSCO	Returns	$\frac{min}{vol}$	Figure 10	Figure 11
	Log-ret	$\frac{op}{max}$	Figure 12	Figure 13

Table 4. *Emef* for fitted *Gh* distributions to AXP and CSCO companies.

Comp	Nature	Values	Gh estimated parameters					Fit.Dist	Figures
			$\hat{\lambda}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\delta}$	$\hat{\mu}$		
AXP	Returns	<i>op.</i>	-1.278	0.01186	0.0118	0.0766	1.005	<i>t</i> -stud	Fig. 7
		<i>min.</i>	-1.247	0.0148	-0.0148	0.0768	1.005	<i>t</i> -stud	
	Log-ret.	<i>max.</i>	-0.5	8.03	-1.37	0.051	0.0105	NIG	Fig. 9
		<i>cl.</i>	-0.5	7.6	-1.24	0.052	0.0103	NIG	
CSCO	Returns	<i>min.</i>	-1.24	0.0148	-0.0148	0.0768	1	<i>t</i> -stud.	Fig. 11
		<i>vol.</i>	-3.82	4.22	4.22	0.613	0.753	<i>t</i> -stud.	
	Log-ret.	<i>op.</i>	-1.26	0.83	-0.83	0.07	0	<i>t</i> -stud.	Fig. 13
		<i>max.</i>	-1.32	0.85	-0.85	0.076	0	<i>t</i> -stud.	

Commentaries.

In view of FIGURE 4 and FIGURE 7 we can say that *t*-student distribution fits well opening and minimum values return for the American Express company AXP, whereas *NIG* distribution fits well maximum and closing log-returns values for the Cysco System company CSCO in view of FIGURE 5 and FIGURE 9.

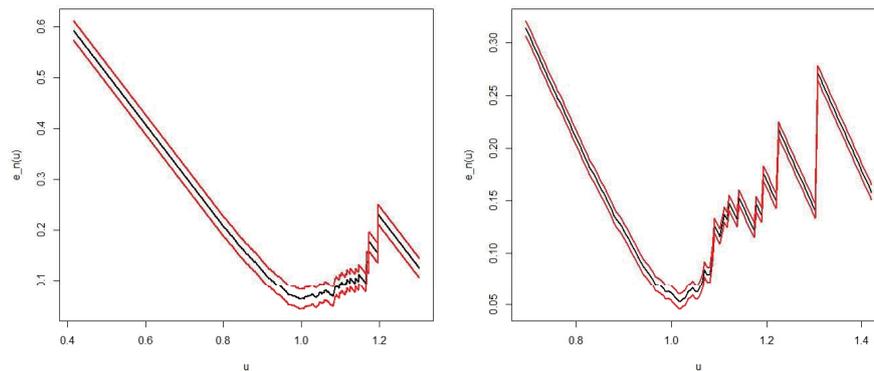


Figure 4. The *emef* for two *t*-Student distribution. The left panel concerns the one with the parameters $\lambda = -1.278$, $\alpha = 0.01186$, $\beta = 0.01186$, $\delta = 0.0766$, $\mu = 1.005$ and the right panel concerns the one with the parameters $\lambda = -1.247$, $\alpha = 0.0148$, $\beta = -0.0148$, $\delta = 0.07683$, $\mu = 1.005$.

6. Conclusion

In this paper we have established almost sure consistency bands for the mean residual life function by using a functional empirical processes approach. Then, we applied these bands for fitting *Gh* distributions to Dow Jones financial data. It is a known fact that these kinds of distributions fit well financial data since they generalize a major part of the classical distribution functions.

We found that Student and *NIG* distributions are good candidates for fitting returns and log-returns data showing their semi-heavy tails.

Acknowledgements

The authors are grateful to the referees and the editor for their careful reading of this paper, for their helpful and constructive criticism that resulted in an improved exposition.

References

- Alexander J. McNeil, Rüdiger, F., & Embrechts, P. (2005). *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton University Press, Princeton.
- Billingsley, P. (1968). *Convergence of Probability measures*, 1st Edition. John Wiley, New-York.
- Csörgő M., & Zitikis, R. (1996). Mean residual life processes. *Ann. Statist.*, 24(4), 1717-1739. <http://dx.doi.org/10.1214/aos/1032298292>.
- Deheuvels, P., & Mason, D. M (2004). General asymptotic confidence bands based on Kernel-type function estimators, *Statistical inference for stochastic processes*, 7(3), 225-277. <http://dx.doi.org/10.1023/B:SISP.0000049092.55534.af>.
- Einmahl, U., & Mason, D. M. (2000). An empirical process approach to the uniform consistency of kernel-type function estimators. *J. Theoret. Probab.*, 13(1), 1-37. <http://dx.doi.org/10.1023/A:1007769924157>.
- Guess, F., & Proschan, F., (1988). Mean residual life : Theory and Applications. In : Rao Krishnaiah, P.R., Rao, C.R. (Eds.), *Handbook of Statistics*, 7, pp. 215-224, North-Holland, Amsterdam.
- Grace, L. Y. (1978). Estimation Of biomedical function. *The Annals of Statistics*, 6(1), 112-116. <http://dx.doi.org/10.1214/aos/1176344070>.
- Hall, W. J., & Wellner, J. A. (1981). Mean residual life. *Statistics and related topics*. Csorgo, M., Dawson, D. A., Rao, J. N. K., and Saleh, A. K. Md. E. (eds.), pp. 169-184. North-Holland, Amsterdam.

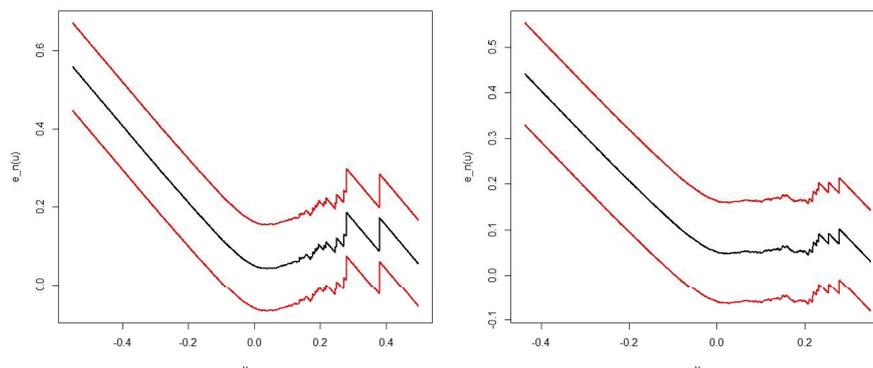


Figure 5. The *emef* for two Normal Inverse Gaussian distributions. The left panel concerns the one with the parameters $\lambda = -0.5$, $\alpha = 8.03$, $\beta = -1.37$, $\delta = 0.051$, $\mu = 0.0105$ and the right panel concerns the one with the parameters $\lambda = -0.5$, $\alpha = 7.6$, $\beta = -1.24$, $\delta = 0.052$, $\mu = 0.0103$.

Kotz, S., & Shanbhag, D. N. (1980). Some new approaches to probability distributions. *Advan. Appl. Probab.*, 12, 903-921. <http://www.jstor.org/stable/1426748>. <http://dx.doi.org/10.2307/1426748>.

Krishnaiah, P. R., & Rao, C. R. (1988). *Quality control and reliability*. In: *Handbook of Statistics*, 7, Amsterdam, North-Holland.

Lo, G. S. (2014). A remark on the asymptotic tightness in $\ell^{+\infty}([a, b])$. arxiv.org/pdf/1405.6342.

Proschan, F., & Serfling, R. J. (1974). *Reliability and Biometry: Statistical Analysis of Life Length: papers*. Proceeding Series, [url=https://books.google.sn/books](https://books.google.sn/books). Philadelphia, PA: SIAM.

Talagrand, M. (1994). Sharper bounds for gaussian and empirical processes, *The Annals of Probability*, 22(1), 28-76. <http://dx.doi.org/10.1214/aop/1176988847>, <http://projecteuclid.org/euclid.aop/1176988847>.

Van Der Vaart, A. W. & Wellner, J. A. (1996). *Weak convergence and empirical processes with applications in statistics*, ISBN 978-1-4757-2547-6, Springer-Verlag New York. <http://dx.doi.org/10.1007/978-1-4757-2545-2>

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).

7. Appendix

7.1 Moment computations

Let Z_1, \dots, Z_n , n i.i.d centered random variables defined on the same probability space with common variance $\mathbb{E}(Z_i^2) = \kappa_1$ and common fourth moment $\mathbb{E}(Z_i^4) = \kappa_2 > 0$. We have

$$\begin{aligned} \mathbb{E}[T(n, u, \delta)]^4 &= \mathbb{E}(Z_1 + Z_2 + \dots + Z_n)^4 = \mathbb{E}\left(\sum_{k=1}^n Z_k^4 + 6 \sum_{1 \leq i < j \leq n} Z_i^2 Z_j^2\right) \\ &= \sum_{k=1}^n \mathbb{E}(Z_k^4) + 6 \sum_{1 \leq i < j \leq n} \mathbb{E}(Z_i^2 Z_j^2) \end{aligned} \tag{14}$$

since, for distinct i, j, k and l ,

$$\mathbb{E}(Z_i Z_j^3) = \mathbb{E}(Z_i Z_j^2 Z_k) = \mathbb{E}(Z_i Z_j Z_k Z_l) = 0,$$

by using independence plus the fact that $\mathbb{E}(Z_i) = 0$. Using independence again,

$$\mathbb{E}(Z_i^2 Z_j^2) = \mathbb{E}(Z_i^2) \mathbb{E}(Z_j^2) = \kappa_1^2 \text{ for } i \neq j.$$

We obtain $\sum_{1 \leq i < j \leq n} \mathbb{E}(Z_i^2 Z_j^2) = \frac{n(n-1)}{2} \kappa_1^2$ since the number of possible couples (i, j) of integers such that $1 \leq i < j \leq n$, is $\binom{n}{2} = \frac{n(n-1)}{2}$. Hence from (14), we deduce that $\mathbb{E}[T(n, u, \delta)]^4 = n\kappa_2 + 3n(n-1)\kappa_1^2$. □

7.2 Proofs of the uniform asymptotic consistency bounds .

We begin to recall the Talagrand bounds and a device of Einmahl and Mason and in SUBSECTION 7.2.2, we give an application of these results.

7.2.1 Talagrand bounds

Before going any further, we recall that a class of measurable real valued functions \mathcal{F} is said to be a *pointwise measurable class* if there exists a countable subclass \mathcal{F}_0 of \mathcal{F} such as, for any function f in \mathcal{F} , we can find a sequence of functions $\{f_m\}_{m \geq 0}$ in \mathcal{F}_0 for which $f_m(x) \rightarrow f(x)$ as $m \rightarrow \infty, x \in \mathbb{R}$. (See Example 2.3.4 in Van Der Vaart, 1996).

Further, let ξ_1, ξ_2, \dots be a sequence of independent Rademacher random variables independent of X_1, X_2, \dots , and \mathbb{G}_m be the functional empirical process indexed by the class of functions \mathcal{F} . The following inequality is essentially due to Talagrand (1994)(see Deheuvels and Mason, 2004 page 269)

Inequality. Let \mathcal{F} be a *pointwise measurable class* of functions satisfying for some $0 < M < \infty, \|f\|_\infty \leq M, f \in \mathcal{F}, .$
Then for all $t > 0$ we have,

$$\begin{aligned} \mathbb{P}\left\{ \max_{1 \leq m \leq n} \|\sqrt{m} \mathbb{G}_m\|_{\mathcal{F}} \geq A_1 \left(\mathbb{E} \left\| \sum_{i=1}^n \xi_i f(X_i) \right\|_{\mathcal{F}} + t \right) \right\} \\ \leq 2 \left(\exp(-A_2 t^2 / n \sigma_{\mathcal{F}}^2) + \exp(-A_2 t / M) \right), \end{aligned} \tag{15}$$

where $\sigma_{\mathcal{F}}^2 = \sup_{f \in \mathcal{F}} \text{Var}(f(X))$ and A_1, A_2 are universal constants.

And the lemma below of Einmahl and Mason (2000) is very helpful for obtaining bounds on this quantity, when the class \mathcal{F} has a polynomial covering number.

Assume that there exists a finite valued measurable function G , called an envelope function, which satisfies for all $x \in \mathbb{R}, G(x) \geq \sup_{f \in \mathcal{F}} |f(x)|$. We define for $0 < \epsilon < 1$

$$N(\epsilon, \mathcal{F}) := \sup_Q N\left(\epsilon \sqrt{Q(G^2)}, \mathcal{F}, d_Q\right)$$

where the supremum is taken over all probability measures Q on \mathbb{R} for which $0 < Q(G^2) := \int G^2(y)Q(dy) < \infty$ and d_Q is the $L_2(Q)$ -metric. As usual, $N(\epsilon, \mathcal{F}, d_Q)$ is the minimal number of balls $\{g : d_Q(g, f) < \epsilon\}$ of d_Q -radius ϵ needed to cover \mathcal{F} . Here is the device of Einmahl and Mason (2000).

Lemma 3. (Einmahl - Mason 2000) Let \mathcal{F} be a pointwise measurable class of bounded functions such that for some constants $\beta > 0, \nu > 0, C > 1, \sigma \leq 1/(8C)$ and function G as above, the following four conditions hold:

(A.1) $\mathbb{E}[G^2(X)] \leq \beta^2;$

(A.2) $N(\epsilon, \mathcal{F}) \leq C\epsilon^{-\nu}, 0 < \epsilon < 1;$

(A.3) $\sigma_0^2 := \sup_{f \in \mathcal{F}} \mathbb{E}[f^2(X)] \leq \sigma^2;$

(A.4) $\sup_{f \in \mathcal{F}} \|f\|_\infty \leq \frac{1}{2\sqrt{\nu+1}} \sqrt{n\sigma^2 / \log(\beta \vee 1/\sigma)}.$

Then we have for some absolute constant $A,$

$$\mathbb{E} \left\| \sum_{i=1}^n \xi_i f(X_i) \right\|_{\mathcal{F}} \leq A \sqrt{vn\sigma^2 \log(\beta \vee 1/\sigma)}. \tag{16}$$

7.2.2 APPLICATION

Put $\ell_u(x) = \ell(x)\mathbb{I}_{(x>u)},$ with $\ell(x) = 1$ or $\ell(x) = x,$ and let $\mathcal{F} = \{\ell_u, u \in I\}.$ \mathcal{F} is pointwise measurable since it suffices to take $\mathcal{F}_0 = \{\ell_u, u \in I \cap \mathbb{Q}\},$ where \mathbb{Q} is the set of irrationnal numbers.

Next $G = \max(|\ell(u_0)|, |\ell(u_1)|) = M > 0$ is an envelope of \mathcal{F} since we have

$$\sup_{u \in I} |\ell_u(x)| \leq |\ell(x)| \leq \max(|\ell(u_0)|, |\ell(u_1)|), \quad \forall u_0 \leq x \leq u_1.$$

Remark that if $\ell(x) = 1$ then $M = 1$ and if $\ell(x) = x$ then $M = \max(|u_0|, |u_1|).$ We have $\sigma_{\mathcal{F}}^2 = \sup_{f \in \mathcal{F}} \text{Var}(f(X)) \leq M^2.$ So

we may use Talagrand’s inequality. It remains to check points of **Lemma 3.** Points (A.1) and (A.3) are obvious with $\beta = M = \sigma.$ To check (A.2), consider any probability Q on $\mathbb{R}.$ We get for $(u, v) \in I^2, u \leq v,$

$$d_Q^2(\ell_u, \ell_v) = \int (\ell_u - \ell_v)^2(x) dQ(x) \leq M^2 Q([u, v]). \tag{17}$$

By a classical result in probability in $\mathbb{R},$ for any given $0 < \epsilon < 1,$ we may cover $[u_0, u_1]$ by at most $m = \lceil \frac{u_1 - u_0}{\epsilon} \rceil$ sub-intervals $[s_{i-1}, s_i]$ such that $Q([s_{i-1}, s_i]) < \epsilon^2, i = 1, \dots, m.$ ($\lceil x \rceil$ stands for the smallest positive integer greater than or equal to x). Let $C = (m + 1)\epsilon,$ we have $m < C\epsilon^{-1}.$ For any $u \in [u_0, u_1],$ there exists $i \in \{1, \dots, m\}$ such as $s_{i-1} \leq u \leq s_i$ with $Q([u, s_i]) < \epsilon^2,$ so the corresponding $\ell_u \in \mathcal{F}$ is such that

$$d_Q(\ell_u, \ell_{s_i}) < \epsilon M = \epsilon \sqrt{Q(G^2)} \quad \text{from (17)}.$$

To finish $m = N(\epsilon \sqrt{Q(G^2)}, \mathcal{F}, d_Q) < C\epsilon^{-1}$ and

$$N(\epsilon, \mathcal{F}) = \sup_Q N(\epsilon \sqrt{Q(G^2)}, \mathcal{F}, d_Q) \leq C\epsilon^{-1}.$$

Now we take $\beta^2 = \sigma^2 = \max(2, \max(|\ell(u_0)|, |\ell(u_1)|)) = M_1.$ Finally for

$$n \geq \frac{8M^2 \log M_1}{M_1^2},$$

we have

$$\mathbb{E} \left\| \sum_{i=1}^n \xi_i g(X_i) \right\|_{\mathcal{F}} \leq C_{\mathcal{F}} \sqrt{n}, \tag{18}$$

where $C_{\mathcal{F}} = A M_1 \sqrt{\log M_1},$ since all the points of the **Lemma 3** are checked. Now we are going to apply the inequality (15) first for the class of functions

$$\mathcal{F}_1 = \{\ell_u(x) = g_u(x), u \in I\}.$$

In this case, $M_1 = 2$ since $\ell(x) = 1$, for any $u_0 \leq x \leq u_1$, and

$$\mathbb{E} \left\| \sum_{i=1}^n \xi_i g_u(X_i) \right\|_{\mathcal{F}_1} = D_{n,1} \leq C_{\mathcal{F}_1} \sqrt{n}, \text{ where } C_{\mathcal{F}_1} = 2A \sqrt{\log 2}.$$

Let $\varepsilon > 0$, $n_1 \geq 2 \log 2$ and t_0 such that

$$\exp\left(\frac{-A_2 t_0^2}{n_1}\right) \leq \frac{\varepsilon}{8}, \text{ and } \exp(-A_2 t_0) \leq \frac{\varepsilon}{8} \text{ and } t_0 < \sqrt{n_1}.$$

(Remind that $\sigma_{\mathcal{F}}^2 = 1$.)

Then

$$\mathbb{P}\left\{ \max_{1 \leq m \leq n} \|\sqrt{m} G_m\|_{\mathcal{F}_1} \geq A_1 \left(\mathbb{E} \left\| \sum_{i=1}^n \xi_i g_u(X_i) \right\|_{\mathcal{F}_1} + t_0 \right) \right\} \leq \varepsilon/2.$$

So for $n \geq n_1$, we arrive at

$$\mathbb{P}\left(|\mathbb{P}_n(g_u) - \mathbb{P}_X(g_u)| < \frac{A_1(D_{n,1} + t_0)}{n}, u \in I \right) > 1 - \varepsilon/2.$$

As $t_0/\sqrt{n} < \sqrt{n_1}/\sqrt{n} \leq 1$, we obtain

$$\frac{A_1(D_{n,1} + t_0)}{n} \leq \frac{A_1 C_{\mathcal{F}_1} \sqrt{n} + A_1 \sqrt{n}}{n} = \frac{A_1 C_{\mathcal{F}_1} + A_1}{\sqrt{n}} = \frac{D_1}{\sqrt{n}},$$

thus

$$\mathbb{P}\left(|\mathbb{P}_n(g_u) - \mathbb{P}_X(g_u)| < \frac{D_1}{\sqrt{n}}, u \in I \right) > 1 - \varepsilon/2 \tag{19}$$

where $D_1 = 2A A_1 \sqrt{\log 2} + A_1$.

Let us use the same method, for the class of functions

$$\mathcal{F}_2 = \{\ell_u(x) = f_u(x), u \in I\}.$$

In this case, $M_1 = \max(2, \max(|u_0|, |u_1|))$ since $\ell(x) = x$, for any $u_0 \leq x \leq u_1$, and

$$\mathbb{E} \left\| \sum_{i=1}^n \xi_i f_u(X_i) \right\|_{\mathcal{F}_2} = D_{n,2} \leq C_{\mathcal{F}_2} \sqrt{n}, \text{ where } C_{\mathcal{F}_2} = A M_1 \sqrt{\log M_1}.$$

Let $n_2 \geq \frac{8M^2 \log M_1}{M_1^2}$ and t_0 such that

$$\exp\left(\frac{-A_2 t_0^2}{n_2}\right) \leq \frac{\varepsilon}{8} \text{ and } \exp(-A_2 t_0) \leq \frac{\varepsilon}{8} \text{ and } t_0 < \sqrt{n_2}.$$

Then

$$\mathbb{P}\left\{ \max_{1 \leq m \leq n} \|\sqrt{m} G_m\|_{\mathcal{F}_2} \geq A_1 \left(\mathbb{E} \left\| \sum_{i=1}^n \xi_i f_u(X_i) \right\|_{\mathcal{F}_2} + t_0 \right) \right\} \leq \varepsilon/2.$$

So for $n \geq n_2$, we deduce that

$$\mathbb{P}\left(|\mathbb{P}_n(f_u) - \mathbb{P}_X(f_u)| < \frac{A_1(D_{n,2} + t_0)}{n}, u \in I \right) > 1 - \varepsilon/2.$$

As $t_0/\sqrt{n} < \sqrt{n_2}/\sqrt{n} \leq 1$, we obtain

$$\frac{A_1(D_{n,2} + t_0)}{n} \leq \frac{A_1 C_{\mathcal{F}_2} \sqrt{n} + A_1 \sqrt{n}}{n} = \frac{A_1 C_{\mathcal{F}_2} + A_1}{\sqrt{n}} = \frac{D_2}{\sqrt{n}},$$

thus

$$\mathbb{P}\left(|\mathbb{P}_n(f_u) - \mathbb{P}_X(f_u)| < \frac{D_2}{\sqrt{n}}, u \in I\right) > 1 - \varepsilon/2, \tag{20}$$

where $D_2 = A A_1 M_1 \sqrt{\log M_1} + A_1$.

Now we use again (5) :

$$|e_n(u) - e(u)| \leq |\mathbb{P}_n(f_u) - \mathbb{P}_X(f_u)| \times |\mathbb{P}_n(g_u)|^{-1} + |\mathbb{P}_X(f_u)| \times \frac{|\mathbb{P}_n(g_u) - \mathbb{P}_X(g_u)|}{|\mathbb{P}_n(g_u)\mathbb{P}_X(g_u)|}.$$

For $u_0 \leq u \leq u_1$, we get

$$0 < \bar{F}(u_1) - \frac{D_1}{\sqrt{n}} \leq \mathbb{P}_X(g_u) - \frac{D_1}{\sqrt{n}} < \mathbb{P}_n(g_u) < \mathbb{P}_X(g_u) + \frac{D_1}{\sqrt{n}}$$

with a probability greater than (w.p.g.t) $1 - \varepsilon/2$ and thus

$$|\mathbb{P}_n(g_u)|^{-1} < \left(\bar{F}(u_1) - \frac{D_1}{\sqrt{n}}\right)^{-1} \text{ w.p.g.t } 1 - \varepsilon/2$$

so we obtain

$$\begin{aligned} &\mathbb{P}\left(|\mathbb{P}_n(f_u) - \mathbb{P}_X(f_u)| \times |\mathbb{P}_n(g_u)|^{-1} < \frac{D_2}{\sqrt{n}} \times \left(\bar{F}(u_1) - \frac{D_1}{\sqrt{n}}\right)^{-1}, u \in I\right) \\ &\geq \mathbb{P}\left(|\mathbb{P}_n(f_u) - \mathbb{P}_X(f_u)| < \frac{D_2}{\sqrt{n}}, u \in I\right) > 1 - \varepsilon/2, \end{aligned}$$

thus

$$\mathbb{P}\left(|\mathbb{P}_n(f_u) - \mathbb{P}_X(f_u)| \times |\mathbb{P}_n(g_u)|^{-1} < \frac{D_2}{\sqrt{n}} \times \left(\bar{F}(u_1) - \frac{D_1}{\sqrt{n}}\right)^{-1}, u \in I\right) > 1 - \varepsilon/2. \tag{21}$$

From the following inequalities :

$$\begin{cases} |\mathbb{P}_n(g_u)|^{-1} \leq \left(\bar{F}(u_1) - \frac{D_1}{\sqrt{n}}\right)^{-1}, \text{ w.p.g.t } 1 - \varepsilon/2, \\ |\mathbb{P}_X(g_u)|^{-1} \leq \bar{F}(u_1)^{-1}, \\ |\mathbb{P}_X(f_u)| \leq \mathbb{E}|X|, \end{cases}$$

and by the same manner, we obtain

$$\begin{aligned} &\mathbb{P}\left(|\mathbb{P}_X(f_u)| \times \frac{|\mathbb{P}_n(g_u) - \mathbb{P}_X(g_u)|}{|\mathbb{P}_n(g_u)\mathbb{P}_X(g_u)|} < \frac{D_1}{\sqrt{n}} \times \mathbb{E}|X| \times \left(\bar{F}(u_1)\bar{F}(u_1) - \frac{D_1}{\sqrt{n}}\right)^{-1}, u \in I\right) \\ &\geq \mathbb{P}\left(|\mathbb{P}_n(g_u) - \mathbb{P}_X(g_u)| < \frac{D_1}{\sqrt{n}}, u \in I\right) > 1 - \varepsilon/2, \end{aligned}$$

thus

$$\mathbb{P}\left(|\mathbb{P}_X(f_u)| \times \frac{|\mathbb{P}_n(g_u) - \mathbb{P}_X(g_u)|}{|\mathbb{P}_n(g_u)\mathbb{P}_X(g_u)|} < \frac{D_1}{\sqrt{n}} \times \mathbb{E}|X| \times \left(\bar{F}(u_1)\bar{F}(u_1) - \frac{D_1}{\sqrt{n}}\right)^{-1}\right) > 1 - \varepsilon/2 \tag{22}$$

By combining (21) and (22), we obtain

$$\begin{aligned} &\mathbb{P}\left(|\mathbb{P}_n(g_u)|^{-1} \times |\mathbb{P}_n(f_u) - \mathbb{P}_X(f_u)| \geq \frac{D_2}{\sqrt{n}} \times \left(\bar{F}(u_1) - \frac{D_1}{\sqrt{n}}\right)^{-1}, u \in I\right) \\ &+ \mathbb{P}\left(|\mathbb{P}_X(f_u)| \times \frac{|\mathbb{P}_n(g_u) - \mathbb{P}_X(g_u)|}{|\mathbb{P}_n(g_u)\mathbb{P}_X(g_u)|} \geq \frac{D_1}{\sqrt{n}} \times \mathbb{E}|X| \times \left(\bar{F}(u_1)\bar{F}(u_1) - \frac{D_1}{\sqrt{n}}\right)^{-1}\right) \\ &\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \leq \varepsilon. \end{aligned}$$

This gives

$$\mathbb{P}\left(|e_n(u) - e(u)| \geq \frac{E_n}{\sqrt{n}}\right) \leq \varepsilon \text{ where } E_n = \frac{1}{\bar{F}(u_1) - D_1/\sqrt{n}}\left(D_2 + \frac{D_1 \times \mathbb{E}|X|}{\bar{F}(u_1)}\right).$$

Finally we conclude by :

$$\mathbb{P}\left(e_n(u) - \frac{E_n}{\sqrt{n}} < e(u) < e_n(u) + \frac{E_n}{\sqrt{n}}, u \in I\right) > 1 - \varepsilon$$

for any $\varepsilon > 0$, any $n \geq n_0$ with

$$E_n = \frac{1}{\bar{F}(u_1) - D_1/\sqrt{n}} \left(D_2 + \frac{D_1 \times \mathbb{E}|X|}{\bar{F}(u_1)} \right) \text{ and } \begin{cases} D_1 = 2AA_1 \sqrt{\log 2} + A_1 \\ D_2 = A A_1 M_1 \sqrt{\log M_1} + A_1, \\ M_1 = \max(2, \max(|u_0|, |u_1|)) \end{cases} \square$$

7.3 Figures

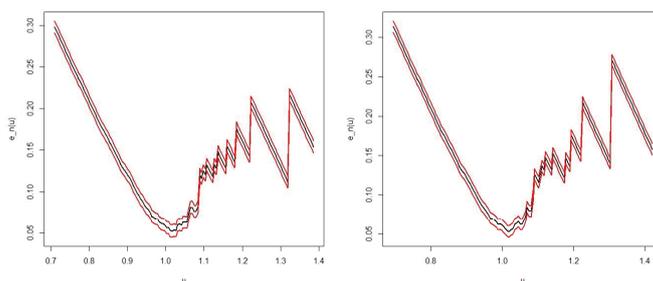


Figure 6. *Emef* for AXP company (monthly data returns). The left panel concerns the opening values and the right one concerns the minimum values.

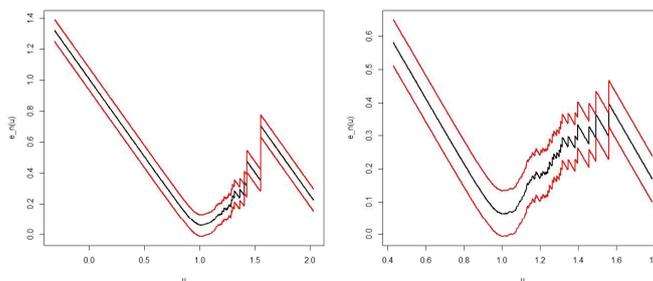


Figure 7. A *t*-Student distribution is fitted to monthly data returns for AXP company (see figure 6). The left panel concerns a *t* distribution with the parameters $\lambda = -1.278, \alpha = 0.01186, \beta = 0.01186, \delta = 0.0766$, and $\mu = 1.005$ fitted to opening values. The right one concerns a *t* distribution with the parameters $\lambda = -1.247, \alpha = 0.0148, \beta = -0.0148, \delta = 0.07683, \mu = 1.005$ fitted to minimum values.

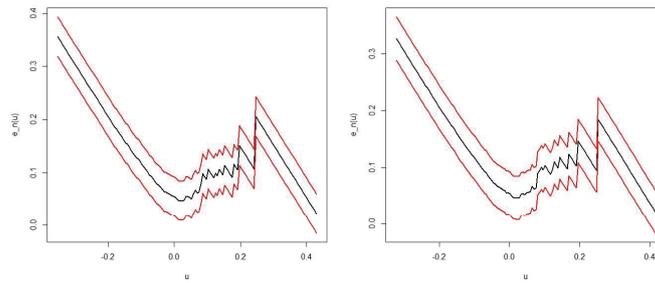


Figure 8. *Emef* for AXP company (monthly data log-returns). The left panel concerns the maximum values and the right panel concerns the closing values.

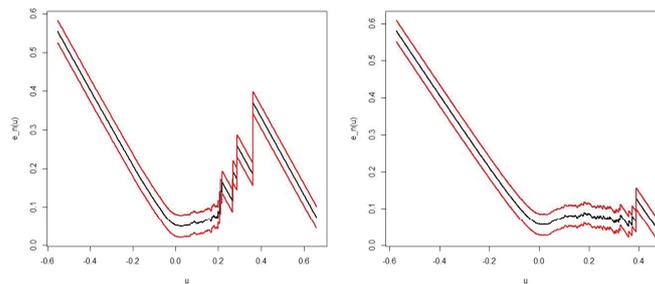


Figure 9. A NIG distribution is fitted to monthly data log-returns for AXP company (see Figure 8). The left panel concerns the one with the parameters $\lambda = -0.5, \alpha = 8.03, \beta = -1.37, \delta = 0.051, \mu = 0.0105$ fitted to maximum values. The right panel concerns the one with the parameters $\lambda = -0.5, \alpha = 7.6, \beta = -1.24, \delta = 0.052, \mu = 0.0103$ fitted to closing values.

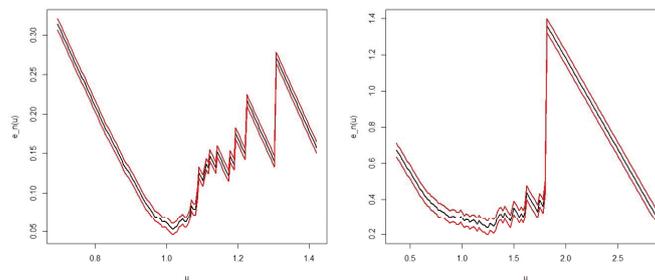


Figure 10. *Emef* for CSCO company (data returns). The left panel concerns monthly minimum values and the right panel concerns monthly volume values.

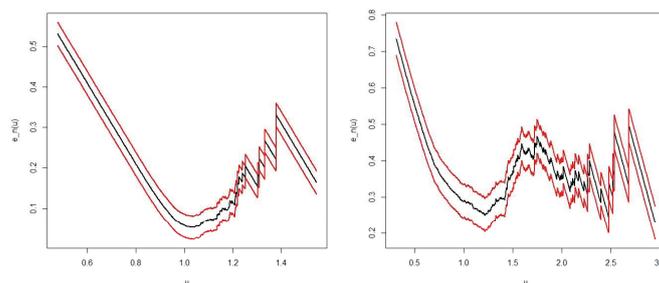


Figure 11. A t -student distribution is fitted to monthly data returns for CSCO company (see Figure 10). The left panel concerns the one with the parameters $\lambda = -1.24, \alpha = 0.014, \beta = -0.014, \delta = 0.076, \mu = 1$ fitted to minimum values. The right panel concerns the one with the parameters $\lambda = -3.82, \alpha = 4.22, \beta = 4.22, \delta = 0.613, \mu = 0.753$ fitted to volume values.

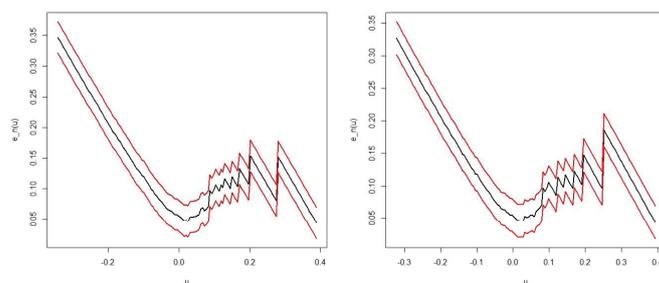


Figure 12. $Emef$ for CSCO company (data log-returns). The left panel concerns monthly opening values and the right panel concerns monthly maximum values.

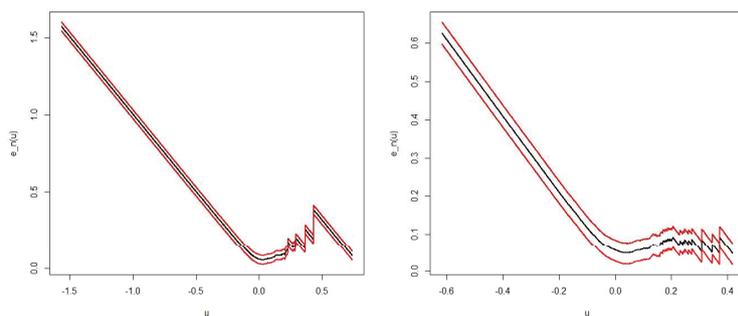


Figure 13. A t -student distribution is fitted to monthly data log-returns for CSCO company (see Figure 12). The left panel concerns the one with the parameters $\lambda = -1.26, \alpha = 0.83, \beta = -0.83, \delta = 0.07, \mu = 0$ fitted to opening values. The right panel concerns the one with the parameters $\lambda = -1.32, \alpha = 0.85, \beta = -0.85, \delta = 0.076, \mu = 0$ fitted to maximum values.