ISSN 1645-6726

Instituto Nacional de Estatística
Statistics Portugal

REVSTAT

Statistical Journal

Volume 5, No.3
November 2007

Catalogação Recomendada

REVSTAT. Lisboa, 2003-
Revstat : statistical journal / ed. Instituto Nacional de Estatística. - Vol. 1, 2003- . - Lisboa I.N.E., 2003- . - 30 cm
Semestral. - Continuação de : Revista de Estatística = ISSN 0873-4275. - edição exclusivamente em inglês
ISSN 1645-6726

CREDITS

- **PUBLISHER**
  - *Instituto Nacional de Estatística (INE)*
    Av. António José de Almeida, 2
    1000-043 LISBOA
    PORTUGAL
    Tel.: +351 218 426 100
    Fax: +351 218 426 364
    Web site: http://www.ine.pt
    Customer Support Service
    (National network): 808 201 808
    (Other networks): +351 226 050 748
- **COVER DESIGN**
  - *Mário Bouçadas, designed on the stain glass window at INE by the painter Abel Manta*
- **LAYOUT AND GRAPHIC DESIGN**
  - *Carlos Perpétuo*
- **PRINTING**
  - *Instituto Nacional de Estatística*
- **EDITION**
  - *350 copies*
- **LEGAL DEPOSIT REGISTRATION**
  - *N.º 191915/03*

PRICE

[VAT 5% included]

- Single issue ………………………………………………… **€ 10**
- Annual subscription (No. 1 Special Issue, No. 2 and No.3)… **€ 24**
- Annual subscription (No. 2, No. 3) ………………………… **€ 16**

Obituary: Radu Theodorescu, 1933–2007

Our dear colleague, Professor Emeritus Radu Theodorescu, passed away in Québec on August 14.

I met Professor Theodorescu for the first time in Lisbon, *circa* 1980. We had a long conversation, and in fact I was astonished and proud that he spent a couple of hours asking details about my thesis, instead of enjoying a visit to Lisbon and its outskirts. I was even more surprised when I found echoes of this conversation on his book with Bertin and Cuculescu, *Unimodality of Probability Measures*.

Our second encounter has been at Funchal, when I organized the 23$^{\text{rd}}$ European Meeting of Statisticians. He came to Funchal with his wife and their two children, and we stayed at the same hotel. Both before and after the meeting we exchanged many e-mails, and during this joyful week at Funchal I could observe his helpful and courteuos attitude towards young statisticians, and his commitment to understand in depth all developments in Statistics. I recommended him to M. Ivette Gomes, who had been invited to be the Editor-in-Chief of REVSTAT, as an appropriate colleague to join the board of editors, an invitation that he gracefully accepted.

Radu published almost 200 papers and monographs, in a broad variety of areas of statistics. In the CV he prepared himself, he proposed a classification of his papers in: Probability theory; concentration functions; unimodality; depths, copulas; stochastic processes (general processes, Markov processes, non-Markovian processes, learning theory, Gaussian reciprocal processes, Lévy processes); statistics (quality control, parametric and nonparametric inference, applications to population dynamics, biology, and geology, stochastic approximation, Linnik laws, estimation, quantile domains); programming, games (stochastic programming, game theory, dynamic programming); control systems, information and communication, automata; numerical analysis; mathematical education, biography, philosophy.

Aside from this huge amount of publications reflecting his many contributions towards the advancement of Probability and Statistics, he served the statistical community preparing more than 600 reviews for *Mathematical Reviews*, *Zentralblatt Math* and other scientific retrieving journals and databases. But above all he has been an outstanding Professor, being awarded the title of Professor Emeritus at University Laval, in 2001.

Radu had been a Ph.D. student of Octav Onicescu. The *Mathematics Genealogy Project* records seven of his Ph.D. students. Christian Genest prepared an obituary for the *IMS Bulletin* **36**, that can be read on-line at

`http://archimede.mat.ulaval.ca/pages/genest/publi/IMS-36-8-16.pdf` ;

read also his interview, in

`http://archimede.mat.ulaval.ca/pages/radutheo/Liaison-Interview.pdf` .

All these can be consulted at

`http://archimede.mat.ulaval.ca/pages/radutheo/` .

I am thankful to Marie-José, Radu's widow, for sending me the material he had himself prepared for his CV and list of publications. M. Ivette Gomes asks me to present the condolences of REVSTAT to her and Radu's children for their sad bereavement, shared by all the other editors.

Dinis Pestana

REVSTAT – Statistical Journal
Volume 5, Number 3, November 2007, 229–247

EXTREMES FOR SOLUTIONS TO STOCHASTIC DIFFERENCE EQUATIONS WITH REGULARLY VARYING TAILS

Authors: Manuel G. Scotto
– Departamento de Matemática, UI&D Matemática e Aplicações, Universidade de Aveiro, Portugal
mscotto@ua.pt

Received: March 2007 Revised: June 2007 Accepted: September 2007

Abstract:

- The main purpose of this paper is to look at the extremal properties of

$$X_k = \sum_{j=1}^{\infty} \left(\prod_{s=1}^{j-1} A_{k-s} \right) B_{k-j} , \qquad k \in \mathbb{Z} ,$$

where $(A_k, B_k)_{k\in\mathbb{Z}}$ is a periodic sequence of independent $\mathbb{R}_+^2$-valued random pairs. The so-called complete convergence theorem we prove enable us to give in detail the weak limiting behavior of various functional of the underlying process including the asymptotic distribution of upper and lower order statistics. In particular, we investigate the limiting distribution of the maximum and its corresponding extremal index. An application to a particular class of bilinear processes is included. These results generalize the ones obtained for the stationary case.

Key-Words:

- *periodic stochastic difference equations; extremal index; point processes.*

AMS Subject Classification:

- 62-02; 60G70.

1. INTRODUCTION

A general approach to look at the extremal properties of non-linear processes is through the analysis of stochastic difference equations (SDEs hereafter) of the form

$$\mathbf{X}_k = \mathbf{A}_k \mathbf{X}_{k-1} + \mathbf{B}_k \ , \qquad k \in \mathbb{Z} \ , \tag{1.1}$$

where $(\mathbf{A}_k)$ are $d{\times}d$ random matrices with possibly negative entries, $(\mathbf{B}_k)$ are $[0,\infty)^d$-valued random (column) vectors such that $(\mathbf{A}_k, \mathbf{B}_k)$ are independent and identically distributed (i.i.d.), and independent of the random column vector $\mathbf{X}_0 \in [0,\infty)^d$. The literature of SDEs is vast mainly for i.i.d. and stationary ergodic sequences $(\mathbf{A}_k, \mathbf{B}_k)$. The existence of a solution to (1.1) has been addressed by Kesten [23], Vervaat [37], and Goldie [18]; for more general results see also Brandt *et al.* [8], Bougerol and Picard [7], and Babillot *et al.* [2]. SDEs play a central role in fields such as finance, economics, insurance mathematics and biology. Examples can be found in Dufresne [12], Embrechts *et al.* [13], Baxendale and Khasminskii [6], Stărică [36], Mikosch [28], and Konstantinides and Mikosch [24]. The interest in these equations is generally justified by the fact that many non-linear processes, including (G)ARCH, threshold, and bilinear processes can be embedded in SDEs.

Extremal properties of the solution of one-dimensional SDEs were first studied by de Haan *et al*. [11] and then by Perfekt [31]. de Haan *et al*. [11] proved the convergence of the point processes of exceedances to a compound Poisson process. As an application, these authors obtained the extremal behavior of the ARCH(1) process. Perfekt [31] extended de Haan *et al*.'s results to Markov processes including SDEs as special cases with possibly negative A_k and B_k. More recently, Scotto [35] derived the extremal behavior of stationary solutions of SDEs where $(A_k, B_k)_{k\in\mathbb{Z}}$ are i.i.d. $\mathbb{R}^2_+$-valued random pairs, the distribution of B_1 being heavy-tailed and the distribution of A_1 having relatively lighter tails compared to the one of B_1 (cf. Grincevičius, [20] and Grey, [19]).

The primary objective of this paper is to derive the extremal properties of one-dimensional SDEs when $(X_k)_{k\in\mathbb{Z}}$ forms a periodic sequence, i.e., when there exists an integer $M \geq 1$ such that for every choice of integers $i_1, ..., i_n$, $(X_{i_1}, ..., X_{i_n})$ and $(X_{i_1+M}, ..., X_{i_n+M})$ are identically distributed. We will refer to such a sequence as an M-periodic sequence if M is the smallest integer as above. Note that if $M = 1$ then $(X_k)_{k\in\mathbb{Z}}$ is a stationary sequence. The study of the extremal properties of non-stationary (periodic) stochastic processes plays a central role when modelling environmental time series, because of its wide applicability to the analysis of phenomena such as extreme concentration of air pollution, floods, wind storms, and extreme temperatures. Extreme value theory of non-stationary processes has been discussed under certain conditions. Horowitz [21] considered the model $\log(Y_k) = g(k) + X_k$, for daily ozone maxima $(Y_k)_{k\in\mathbb{Z}}$, where $g(k)$ is a deterministic function and $(X_k)_{k\in\mathbb{Z}}$ is a normal stationary autoregressive process. Ballerini and McCormick [4] discussed limit theory for non-stationary random sequences of the form $Y_k = g(k) + h(k)\,X_k$, where $(X_k)_{k\in\mathbb{Z}}$ is a stationary random sequence, satisfying some mixing conditions, and $h(k)$ is a positive, periodic function with integer period $p > 1$ as the variance function. The authors derived the limiting distribution of the maximum term based on the assumption that the distribution of X_k belongs to the domain of attraction of an extreme value distribution. The results were applied in a rainfall study; see also Ballerini and Waylen [5] and Ballerini [3]. Niu [29] introduced a class of nonlinear additive time series models for daily maxima of ozone concentrations in which both mean levels and variances are nonlinear functions of relevant meteorological variables. As an alternative approach to analyze tropospheric ozone data Niu [30] focus on estimating probabilities of monthly maximum ozone observations exceeding some specific levels, calculating the mean rate of exceedances of daily maximum ozone over the national standard level 120 ppb (parts per billion). For further examples see Coles [9].

Extreme value theory for periodic sequences was first considered by Alpuim [1] who showed that under Leadbetter's D condition (Leadbetter *et al.*, [26]) the only possible limit laws for the normalized maxima of the periodic sequence are the three extreme value distributions. Extensions for randomly indexed periodic sequences under long range dependence conditions were established by Ferreira [16]. Further results can be found in Ferreira [15] who studied the extremal behavior of periodic sequences under local mixing conditions. Generalizations under weaker local mixing conditions have been considered by Ferreira and Martins [17]. More recently, Martins and Ferreira [27] derived the expression of the extremal index (and hence the limiting distribution of the maximum) of a periodic moving average sequence driven by heavy-tailed innovations.

The rest of the paper is organized as follows: Section 2 deals with the tail behavior of $X_r,\ r = 1, ..., M$. Section 3 is devoted to a detailed point process analysis of asymptotic properties of the periodic sequence $(X_k)_{k\in\mathbb{Z}}$. In particular we deduce the maximum limiting distribution and the extremal index. Finally, in Section 4 the results are applied to a particular class of bilinear processes.

2. TAIL BEHAVIOR

Let $(A_k, B_k)_{k\in\mathbb{Z}}$ be a one-dimensional M-periodic sequence of independent $\mathbb{R}^2_+$-valued random pairs, such that $\bar{F}_{B_r}(x) = P(B_r > x)$, $r = 1, ..., M$, are regularly varying with tail index $-\alpha$, for some $\alpha > 0$, i.e.,

$$\bar{F}_{B_r}(x) = x^{-\alpha} L_r(x)\ , \qquad r = 1, ..., M\ , \tag{2.1}$$

for some slowly varying functions $L_r\colon \mathbb{R}_+ \to \mathbb{R}_+$ $(r = 1, ..., M)$ at infinity.

We further assume that the tails are equivalent in the sense that

$$(2.2)\qquad \lim_{x\to\infty} \frac{\bar{F}_{B_l}(x)}{\bar{F}_{B_k}(x)} = \gamma_{l,k} \;, \qquad (0<\gamma_{l,k}<\infty) \quad l,k \in \mathbb{Z} \;.$$

Note that $\gamma_{l,k} = \gamma_{l+M,k}$ and $\gamma_{l,k} = \gamma_{l,k+M}$. In addition, we assume that for $r = 1, ..., M$

$$(2.3)\qquad EA_r^\alpha < 1 \,, \qquad \text{and} \qquad EA_r^{\alpha+\delta} < \infty, \quad \text{for some } \; \delta > 0 \;.$$

Note that no further assumptions are needed since the central role in determining the tail behavior of X_r is played by the distributions F_{B_r}. Furthermore, we assume that X_k admits the representation

$$(2.4)\qquad X_k \;=\; \sum_{j=1}^{\infty} \left(\prod_{s=1}^{j-1} A_{k-s} \right) B_{k-j} \;,$$

where we use the convention $\prod_{s=1}^{0} = 1$. This series representation is possible a.s. by virtue of the assumptions on A_k and B_k. Clearly, $(X_k)_{k\in\mathbb{Z}}$ forms an M-periodic sequence and satisfies the SDEs

$$X_k \;=\; A_k X_{k-1} + B_k \;.$$

We start with the analysis of the tail behavior of X_r, $r = 1, ..., M$. In doing so, the following alternative representation of X_r is very useful.

Proposition 2.1. *For the process defined in* (2.4)*, it holds that for* $r = 1, ..., M$

$$X_r \;=\; \sum_{i=0}^{M-1} X_r^{(i)} \;,$$

with

$$X_r^{(i)} \;=\; \sum_{j=1}^{\infty} \left(\prod_{s=1}^{M(j-1)+i} A_{r-s} \right) B_{r-(j-1)M-i-1} \;.$$

We now begin with a series of results designed to understand the tail behavior of $X_r^{(i)}$ as well as sums of these variables. The tail behavior of $X_r^{(i)}$ will be derived in two stages: first we obtain the tail behavior of the approximation $X_{r,m}^{(i)}$, with $m = KM \;\; (K \ge 1)$, defined as

$$X_{r,m}^{(i)} \;=\; \sum_{j=1}^{m} W_{r,i}^{(j)} \;,$$

with

$$W_{r,i}^{(j)} \;=\; \left(\prod_{s=1}^{M(j-1)+i} A_{r-s} \right) B_{r-(j-1)M-i-1} \;;$$

then the results are extended so that the number of summands can be infinite.

Lemma 2.1. *Let $(A_k, B_k)_{k\in\mathbb{Z}}$ be an M-periodic sequence of independent $\mathbb{R}^2_+$-valued random pairs satisfying* (2.1), (2.2), *and* (2.3). *For a fixed value $0 \le i \le M-1$ and $1 \le j \le m$, we have as $x \to \infty$*

$$P\big(W^{(j)}_{r,i} > x\big) \sim \gamma_{r+M-i-1,r} \left(\prod_{s=1}^{i} E(A^\alpha_{r-s})^j\right) \left(\prod_{s=i+1}^{M} E(A^\alpha_{r-s})^{j-1}\right) P\big(B_r > x\big) . \tag{2.5}$$

Furthermore, for all fixed values $1 \le j_1 < j_2 \le m$ and $0 \le i \le M-1$, as $x \to \infty$

$$\frac{P\Big(W^{(j_1)}_{r,i} > x,\, W^{(j_2)}_{r,i} > x\Big)}{P\big(B_r > x\big)} \to 0 \,, \qquad r = 1, ..., M\,. \tag{2.6}$$

Proof: The first statement follows as an application of Breiman's result (cf. Davis and Resnick [10], p. 1197). Let $C_{s,j_h} = \prod_{s=1}^{M(j_h-1)+i} A_{r-s}$, for $h = 1, 2$. In proving (2.6) observe that

$$\begin{aligned}
P\Big(W^{(j_1)}_{r,i} > x,\, W^{(j_2)}_{r,i} > x\Big) \;=\;& \\
&= P\Big(C_{s,j_1} B_{r-(j_1-1)M-i-1} > x,\; C_{s,j_1} C_{s,j_2} B_{r-(j_2-1)M-i-1} > x\Big) \\
&\le P\Big(C_{s,j_1} \le \epsilon,\; C_{s,j_1} B_{r-(j_1-1)M-i-1} > x\Big) \\
&\quad + P\Big(C_{s,j_1} > \epsilon,\; C_{s,j_1} B_{r-(j_1-1)M-i-1} > x,\; C_{s,j_1} C_{s,j_2} B_{r-(j_2-1)M-i-1} > x\Big) \\
&\le P\Big(C_{s,j_1} 1_{[C_{s,j_1} \le \epsilon]}\, B_{r-(j_1-1)M-i-1} > x\Big) \\
&\quad + P\Big(B_{r-(j_1-1)M-i-1} > \frac{x}{\epsilon},\; C_{s,j_2} B_{r-(j_2-1)M-i-1} > \frac{x}{\epsilon}\Big) \,.
\end{aligned}$$

Now, by Breiman's result

$$\begin{aligned}
\limsup_{x\to\infty} \frac{P\Big(C_{s,j_1} 1_{[C_{s,j_1} \le \epsilon]}\, B_{r-(j_1-1)M-i-1} > x\Big)}{P\big(B_r > x\big)} \;&=\; \gamma_{r+M-i-1,r}\, E\Big(C_{s,j_1} 1_{[C_{s,j_1} \le \epsilon]}\Big)^\alpha \\
&\to 0 \,,
\end{aligned}$$

as $\epsilon \to 0$. Moreover,

$$\begin{aligned}
&\frac{P\Big(B_{r-(j_1-1)M-i-1} > \frac{x}{\epsilon},\; C_{s,j_2} B_{r-(j_2-1)M-i-1} > \frac{x}{\epsilon}\Big)}{P\big(B_r > x\big)} \sim \\
&\qquad\qquad \sim\; \epsilon^{2\alpha}\, \gamma_{r+M-i-1,r}\, E(C_{s,j_2})^\alpha\, P\Big(B_{r-(j_2-1)M-i-1} > x\Big) \,,
\end{aligned}$$

as $x \to \infty$. Note that as $\epsilon \to 0$, the right-hand side converges to 0. This completes the proof. □

Lemma 2.2. *Let $(A_k, B_k)_{k\in\mathbb{Z}}$ be an M-periodic sequence of independent $\mathbb{R}^2_+$-valued random pairs satisfying* (2.1), (2.2), *and* (2.3). *For a fixed value of* $0 \le i \le M-1$

$$\lim_{x\to\infty} \frac{P\big(X^{(i)}_{r,m} > x\big)}{P\big(B_r > x\big)} = \left(\prod_{s=1}^{i} EA^\alpha_{r-s}\right) \frac{1-\left(\prod_{s=1}^{M} EA^\alpha_{r-s}\right)^m}{1-\prod_{s=1}^{M} EA^\alpha_{r-s}}\, \gamma_{r+M-i-1,r}\ , \tag{2.7}$$

for $r = 1, ..., M$. *Moreover, as* $m \to \infty$

$$\lim_{x\to\infty} \frac{P\big(X^{(i)}_{r} > x\big)}{P\big(B_r > x\big)} = \frac{\prod_{s=1}^{i} EA^\alpha_{r-s}}{1-\prod_{s=1}^{M} EA^\alpha_{r-s}}\, \gamma_{r+M-i-1,r}\ . \tag{2.8}$$

Proof: The first statement follows as an application of Lemma 2.1 in Davis and Resnick [10] and Lemma 2.1. The proof is complete upon showing that by letting $m \to \infty$ we obtain (2.8). First note that the first statement implies that

$$\begin{aligned}\liminf_{x\to\infty} \frac{P\big(X^{(i)}_{r} > x\big)}{P\big(B_r > x\big)} &\ge \liminf_{x\to\infty} \frac{P\big(X^{(i)}_{r,m} > x\big)}{P\big(B_r > x\big)} \\ &= \left(\prod_{s=1}^{i} EA^\alpha_{r-s}\right) \frac{1-\left(\prod_{s=1}^{M} EA^\alpha_{r-s}\right)^m}{1-\prod_{s=1}^{M} EA^\alpha_{r-s}}\, \gamma_{r+M-i-1,r}\ .\end{aligned}$$

Hence, as $m \to \infty$

$$\liminf_{x\to\infty} \frac{P\big(X^{(i)}_{r} > x\big)}{P\big(B_r > x\big)} \ge \frac{\prod_{s=1}^{i} EA^\alpha_{r-s}}{1-\prod_{s=1}^{M} EA^\alpha_{r-s}}\, \gamma_{r+M-i-1,r}\ .$$

The arguments needed to get the upper bound follow closely the arguments outlined in Resnick ([33], p. 228): decompose the event $\big[X^{(i)}_r > x\big]$ according to whether $\big[\max_{j\in\mathbb{N}} W^{(j)}_{r,i} > x\big]$ or $\big[\max_{j\in\mathbb{N}} W^{(j)}_{r,i} \le x\big]$

$$\begin{aligned} P\big(X^{(i)}_r > x\big) &= P\Big(X^{(i)}_r > x,\ \max_{j\in\mathbb{N}} W^{(j)}_{r,i} > x\Big) + P\Big(X^{(i)}_r > x,\ \max_{j\in\mathbb{N}} W^{(j)}_{r,i} \le x\Big) \\ &\le P\left(\bigcup_{j\in\mathbb{N}} W^{(j)}_{r,i} > x\right) + P\left(\sum_{j=1}^{\infty} W^{(j)}_{r,i}\, 1_{\{W^{(j)}_{r,i}\le x\}} > x,\ \max_{j\in\mathbb{N}} W^{(j)}_{r,i} \le x\right) \\ &\le \sum_{j=1}^{\infty} P\big(W^{(j)}_{r,i} > x\big) + P\left(\sum_{j=1}^{\infty} W^{(j)}_{r,i}\, 1_{\{W^{(j)}_{r,i}\le x\}} > x\right). \end{aligned}$$

By Markov's inequality

$$\begin{aligned} \frac{P\big(X^{(i)}_r > x\big)}{P\big(B_r > x\big)} &\le \frac{\sum_{j=1}^{\infty} P\big(W^{(j)}_{r,i} > x\big)}{P\big(B_r > x\big)} + \frac{\sum_{j=1}^{\infty} EW^{(j)}_{r,i}\, 1_{\{W^{(j)}_{r,i}\le x\}}}{x\, P\big(B_r > x\big)} \\ &= I(x) + J(x)\ . \end{aligned}$$

To handle $I(x)$, note that by Kamarata's Theorem quoted in Resnick ([33], p. 17), the result in (2.5) along with condition (2.3) and dominated convergence, lead us to obtain

$$\lim_{x\to\infty} I(x) \;=\; \frac{\prod_{s=1}^{i} EA_{r-s}^{\alpha}}{1-\prod_{s=1}^{M} EA_{r-s}^{\alpha}}\;\gamma_{r+M-i-1,r}\;.$$

For $J(x)$ let us start by considering the case $0<\alpha<1$. By Lemma 2.1, the distribution tail of $W_{r,i}^{(j)}$ is regularly varying with index $-\alpha$. Now

$$\frac{EW_{r,i}^{(j)}\,1_{\{W_{r,i}^{(j)}\le x\}}}{x\,P\big(B_r>x\big)} \;=\; \frac{EW_{r,i}^{(j)}\,1_{\{W_{r,i}^{(j)}\le x\}}}{x\,P\big(W_{r,i}^{(j)}>x\big)}\,\frac{P\big(W_{r,i}^{(j)}>x\big)}{P\big(B_r>x\big)}\;.$$

From an integration by parts along with the result in (2.5), and Kamarata's Theorem

$$\text{(2.9)}\qquad \frac{EW_{r,i}^{(j)}\,1_{\{W_{r,i}^{(j)}\le x\}}}{x\,P\big(W_{r,i}^{(j)}>x\big)} \;\to\; \alpha(1-\alpha)^{-1}\;, \qquad x\to\infty\;.$$

Since $P(B_r>x)$ is regularly varying with index $-\alpha$ we can use its Kamarata representation and (2.5) to obtain that for sufficiently large x and some constant $K>0$

$$\text{(2.10)}\qquad \frac{P\big(W_{r,i}^{(j)}>x\big)}{P\big(B_r>x\big)} \;\le\; K\,\gamma_{r+M-i-1,r}\left(\prod_{s=1}^{i} E(A_{r-s}^{\alpha})^{j}\right)\left(\prod_{s=i+1}^{M} E(A_{r-s}^{\alpha})^{j-1}\right)\;,$$

for $r=1,...,M$. Combining (2.9) and (2.10), we conclude, for sufficiently large x

$$\frac{EW_{r,i}^{(j)}\,1_{\{W_{r,i}^{(j)}\le x\}}}{x\,P\big(B_r>x\big)} \;\le\; K_1\,\gamma_{r+M-i-1,r}\left(\prod_{s=1}^{i} E(A_{r-s}^{\alpha})^{j}\right)\left(\prod_{s=i+1}^{M} E(A_{r-s}^{\alpha})^{j-1}\right)\;,$$

for some constant $K_1>0$. This bound is summable providing, by dominated convergence

$$\begin{aligned}\limsup_{x\to\infty} J(x) \;&\le\; K_1\sum_{j=1}^{\infty}\gamma_{r+M-i-1,r}\left(\prod_{s=1}^{i} E(A_{r-s}^{\alpha})^{j}\right)\left(\prod_{s=i+1}^{M} E(A_{r-s}^{\alpha})^{j-1}\right)\\ &=\; K_1\,\frac{\prod_{s=1}^{i} EA_{r-s}^{\alpha}}{1-\prod_{s=1}^{M} EA_{r-s}^{\alpha}}\;\gamma_{r+M-i-1,r}\end{aligned}$$

and hence

$$\text{(2.11)}\qquad \limsup_{x\to\infty}\frac{P\big(X_r^{(i)}>x\big)}{P\big(B_r>x\big)} \;\le\; (K_1+1)\,\frac{\prod_{s=1}^{i} EA_{r-s}^{\alpha}}{1-\prod_{s=1}^{M} EA_{r-s}^{\alpha}}\;\gamma_{r+M-i-1,r}\;.$$

If $\alpha \geq 1$, we proceed as follows:

pick $\beta \in (\alpha, \alpha\delta^{-1})$ and consider $A^{(i)} = \sum_{j=1}^{\infty} \prod_{s=1}^{M(j-1)+i} A_{r-s}$, $(i=1,...,M)$ and $P_j^{(i)} = \big(\prod_{s=1}^{M(j-1)+i} A_{r-s}\big)\{A^{(i)}\}^{-1}$, $(i=1,...,M,\ j\in\mathbb{N})$.

By Jensen's inequality

$$\begin{aligned}
\big(X_r^{(i)}\big)^{\beta} &= \{A^{(i)}\}^{\beta}\left(\sum_{j=1}^{\infty} P_j^{(i)} B_{r-(j-1)M-i-1}\right)^{\beta} \\
&\leq \{A^{(i)}\}^{\beta} \sum_{j=1}^{\infty} P_j^{(i)} B^{\beta}_{r-(j-1)M-i-1} \\
&= \{A^{(i)}\}^{\beta-1} \sum_{j=1}^{\infty}\left(\prod_{s=1}^{M(j-1)+i} A_{r-s}\right) B^{\beta}_{r-(j-1)M-i-1} ,
\end{aligned}$$

providing

$$\frac{P\big(X_r^{(i)} > x\big)}{P\big(B_r > x\big)} \leq \frac{P\Big(\{A^{(i)}\}^{\beta-1} \sum_{j=1}^{\infty}\Big(\prod_{s=1}^{M(j-1)+i} A_{r-s}\Big) B^{\beta}_{r-(j-1)M-i-1} > x^{\beta}\Big)}{P\big(B_r^{\beta} > x^{\beta}\big)} .$$

Using the fact that $P(B_r^{\beta} > x) \in RV_{-\alpha\beta^{-1}}$ with $\delta < \alpha\beta^{-1}$, for $r=1,...,M$ and $i = 0,...,M-1$ it follows that

$$\begin{aligned}
\limsup_{x\to\infty} \frac{P\big(X_r^{(i)} > x\big)}{P\big(B_r > x\big)} &\leq (K_1+1) \sum_{j=1}^{\infty}\left(\prod_{s=1}^{i} E(A_{r-s}^{\alpha})^{j}\right)\left(\prod_{s=i+1}^{M} E(A_{r-s}^{\alpha})^{j-1}\right) \\
&\qquad \times \{EA^{(i)}\}^{\alpha(1-\beta^{-1})}\, \gamma_{r+M-i-1,r} < \infty .
\end{aligned} \tag{2.12}$$

On the other hand, for any $\epsilon > 0$

$$\frac{P\big(X_r^{(i)} > x\big)}{P\big(B_r > x\big)} \leq \frac{P\Big(\sum_{j=1}^{m} W_{r,i}^{(j)} > (1-\epsilon)\,x\Big)}{P\big(B_r > x\big)} + \frac{P\Big(\sum_{j=m+1}^{\infty} W_{r,i}^{(j)} > \epsilon\, x\Big)}{P\big(B_r > x\big)} ,$$

and for (2.7) and (2.11)

$$\begin{aligned}
\limsup_{x\to\infty} \frac{P\big(X_r^{(i)} > x\big)}{P\big(B_r > x\big)} &\leq (1-\epsilon)^{-\alpha}\left(\prod_{s=1}^{i} EA_{r-s}^{\alpha}\right)\frac{1-\Big(\prod_{s=1}^{M} EA_{r-s}^{\alpha}\Big)^{m}}{1-\prod_{s=1}^{M} EA_{r-s}^{\alpha}}\, \gamma_{r+M-i-1,r} \\
&\quad + K_1\,\epsilon^{-\alpha} \\
&\quad \times \sum_{j=m+1}^{\infty}\left(\prod_{s=1}^{i} E(A_{r-s}^{\alpha})^{j}\right)\left(\prod_{s=i+1}^{M} E(A_{r-s}^{\alpha})^{j-1}\right)\gamma_{r+M-i-1,r} ,
\end{aligned}$$

for the case $0 \leq \alpha \leq 1$ with a similar bound for the second piece provided by (2.12) when $\alpha \geq 1$. Let $m \to \infty$ and then send $\epsilon \to 0$ to obtain

$$\limsup_{x\to\infty} \frac{P\big(X_r^{(i)} > x\big)}{P\big(B_r > x\big)} \leq \frac{\prod_{s=1}^{i} EA_{r-s}^{\alpha}}{1-\prod_{s=1}^{M} EA_{r-s}^{\alpha}}\, \gamma_{r+M-i-1,r}$$

and this combined with the liminf statement concludes the proof. □

Combining Lemmas 2.1 and 2.2 yields the following result.

Theorem 2.1. *Let $(X_k)_{k\in\mathbb{Z}}$ be the M-periodic sequence defined in* (2.4). *Let $(A_k,B_k)_{k\in\mathbb{Z}}$ be an M-periodic sequence of independent $\mathbb{R}^2_+$-valued random pairs satisfying* (2.1), (2.2), *and* (2.3). *For $r=1,...,M$*

$$\lim_{x\to\infty}\frac{P\big(X_r>x\big)}{P\big(B_r>x\big)} = \frac{1}{1-\prod_{s=1}^{M} EA^{\alpha}_{r-s}}\sum_{i=0}^{M-1}\gamma_{r+M-i-1,r}\left(\prod_{s=1}^{i} EA^{\alpha}_{r-s}\right). \tag{2.13}$$

Proof: Note that by Lemma 2.1 in Davis and Resnick [10] it is sufficient to show that for $0\le i_1<i_2\le M-1$, as $x\to\infty$

$$\frac{P\Big(X_r^{(i_1)}>x,\ X_r^{(i_2)}>x\Big)}{P\big(B_r>x\big)}\sim 0\ , \qquad r=1,...,M\,. \tag{2.14}$$

Now an argument similar to the one in the proof of Lemma 2.1 shows that (2.14) holds. □

3. POINT PROCESS APPROACH

In this section we investigate the limit behavior of a sequence of point processes based on the periodic sequence $(X_k)_{k\in\mathbb{Z}}$. Since our results are based on point process theory, we briefly discuss some notation and background about point processes; for further details see Kallenberg [22] and Resnick [33]. Let $(\Omega,\mathcal{F},\mathcal{P})$ be a probability space and E a state space where points reside and assume that E is Euclidian. Let $\mathcal{E}$ be the σ-algebra on E generated by open sets of E. For $x\in E$, define $\epsilon_x(\cdot)$ on $\mathcal{E}$ as the simple point measure with unit mass at x. Let $\{x_j\}$ be a countable collection of points on E. A point measure N on $\mathcal{E}$ is defined to be

$$N(\cdot) = \sum_{j=1}^{\infty}\epsilon_{x_j}(\cdot)\ ,$$

which is a non-negative integer valued Radon measure on compact subsets of E. Let $M_p(E)$ be the class of such Radon measures on $\mathcal{E}$ and $\mathcal{M}_p(E)$ the smallest σ-algebra, making maps $N\to N(A^*)$ measurable, where $N\in M_p(E)$ and $A^*\in\mathcal{E}$. $\mathcal{M}_p(E)$ can be made into a complete separable metric space, hence we assume that it is a metric space with vague metric d. A point process on E is a measurable map from $(\Omega,\mathcal{F})$ to $\big(M_p(E),\mathcal{M}_p(E)\big)$. Let $C_K^+(E)$ be the set of all continuous, non-negative functions on the state space E with compact support. If $N_n\in M_p(E)$ then N_n converges vaguely to N $(N_n\Rightarrow N)$ if $N_n(f)$ converges to $N(f)$ for every

$f \in C_K^+(E)$, where $N(f) = \int f\, dN$. A Poisson process on $(E, \mathcal{E})$ with mean measure μ is a point process N such that, for every $A^* \in \mathcal{E}$, $N(A^*)$ is a Poisson random variable with mean measure $\mu(A^*)$. If $A_1^*, ..., A_m^*$ are mutually independent sets then $N(A_1^*), ..., N(A_m^*)$ are independent random variables. We call N a Poisson random measure with mean measure μ or PRM(μ) for short.

In this section, we investigate the limiting behavior of a sequence of point processes $(N_n)_{n\in\mathbb{N}}$ defined as

$$N_n = \sum_{k=1}^{\infty} \epsilon_{\{k/n,\, a_n^{-1} X_k\}} \ ,$$

based on $(a_n^{-1} X_k)_{k\in\mathbb{Z}}$ with the sequence of norming constants $(a_n)_{n\in\mathbb{N}}$ satisfying

$$\lim_{n\to\infty} n P\big(B_r > a_n x\big) = \tau_r \ , \qquad (\tau_r > 0), \quad r = 1, ..., M \, .$$

Note that such a sequence exists by the assumption of regular variation of each $\bar{F}_{B_r}$, $(r = 1, ..., M)$, and implies that

$$n P\big(X_r > a_n x\big) \ \to \ \tau_r \left\{ \frac{1}{1 - \prod_{s=1}^{M} E A_{r-s}^{\alpha}} \sum_{i=0}^{M-1} \left(\prod_{s=1}^{i} E A_{r-s}^{\alpha} \right) \gamma_{r+M-i-1,r} \right\} ,$$

as $n \to \infty$. It is important to point out the fact that $\tau_r = \tau_h \gamma_{r,l}$, for $r, l \in \{1, ..., M\}$. Hence, without lost of generality it will be assumed that $\tau_r = \tau_1 \gamma_{r,1}$ with $\tau_1 = x^{-\alpha}$.

The main result of this section is formalized through the following theorem, which discusses the weak convergence of the sequence of point processes $(N_n)_{n\in\mathbb{N}}$ to a function of PRM. For simplicity of notation we define $E_h = (0, \infty) \times [-\infty, \infty]^h \backslash \{\mathbf{0}\}$, with $h \in \mathbb{N}$.

Theorem 3.1. *Let $(X_k)_{k\in\mathbb{Z}}$ be an M-periodic sequence defined as in* (2.4) *where $(A_k, B_k)_{k\in\mathbb{Z}}$ is an M-periodic sequence of independent $\mathbb{R}_+^2$-valued random pairs satisfying* (2.1), (2.2), *and* (2.3). *Then, as $n \to \infty$*

$$N_n = \sum_{k=1}^{\infty} \epsilon_{\left\{\frac{k}{n},\, a_n^{-1} X_k\right\}} \ \Rightarrow \ N = \sum_{r=1}^{M} \sum_{i=0}^{M-1} \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} \epsilon_{\left\{T_{k,r}^{(i)},\, J_{k,r}^{(i)} U_{k,1,r} \cdots U_{k,M(j-1)+i-1,r}\right\}} \ ,$$

in the space $M_p(E_1)$, where $\sum_{k=1}^{\infty} \epsilon_{\left\{T_{k,r}^{(i)},\, J_{k,r}^{(i)}\right\}}$ are PRM $(dt \times d\nu_{r,i})$ with

$$\nu_{r,i} = \frac{1}{M} \, \gamma_{r,1} \, \gamma_{r+M-i-1,r} \, \mu(dx) \ ,$$

where $\mu(dx) = \alpha\, x^{-\alpha-1} 1_{(0,\infty]}(x)\, dx$ and $(U_{k,1,r}, ..., U_{k,M,r})$ having the same distribution as $(A_1, ..., A_M)$.

Proof: First note that

$$\sum_{k=1}^{\infty} \epsilon_{\left\{\frac{k}{n},\, a_n^{-1} X_k\right\}} = \sum_{r=1}^{M} \sum_{k=1}^{\infty} \epsilon_{\left\{\frac{(k-1)M+r}{n},\, a_n^{-1} X_{(k-1)M+r}\right\}} .$$

As an application of Proposition 3.2 in Feigin *et al.* [14], for fixed values of $r = 1..., M$ and $i = 0, ..., M-1$, it follows that

$$\begin{aligned}\sum_{k=1}^{\infty} \epsilon_{\left\{\frac{(k-1)M+r}{n},\; a_n^{-1}\left(B_{k-(j-1)M-i-1)},\, j=1,...,m\right),\; A_{k-s},\, s=1,...,M(j-1)+i\right\}} \Rightarrow& \\ \Rightarrow \sum_{k=1}^{\infty} \epsilon_{\left\{T_{k,r}^{(i)},\, J_{k,r}^{(i)} \mathbf{e}_1, \infty, U_{k,1,r}, ..., U_{kM(j-1)+i,r}\right\}}& \\ + \sum_{k=1}^{\infty} \epsilon_{\left\{T_{k,r}^{(i)},\, J_{k}^{(i)} \mathbf{e}_2, U_{k,1,r}, \infty, ..., U_{k,M(j-1)+i,r}\right\}}& \\ \vdots\qquad& \\ + \sum_{k=1}^{\infty} \epsilon_{\left\{T_{k,r}^{(i)},\, J_{k,r}^{(i)} \mathbf{e}_m, U_{k,1,r}, ..., U_{k,M(j-1)+i,r}, \infty\right\}}& ,\end{aligned}$$

in $M_p\left(E_m \times (0,\infty)^{M(j-1)+i}\right)$, where $\mathbf{e}_s$ is the unit vector in $\mathbb{R}^m$ with 1 in the s-th component and the rest zero. By the lines of reasoning given in Resnick and Van den Berg ([34], Theorem 4.1) it follows that, for a fixed value of $r = 1..., M$ and $i = 0, ..., M-1$

$$\sum_{k=1}^{\infty} \epsilon_{\left\{\frac{(k-1)M+r}{n},\, a_n^{-1} X_{(k-1)M+r}^{(i)}\right\}} \Rightarrow \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} \epsilon_{\left\{T_{k,r}^{(i)},\, J_{k,r}^{(i)} U_{k,1,r} \cdots U_{k,M(j-1)+i,r}\right\}} ,$$

in $M_p(E_1)$. Next we have to show that the point processes

$$N_n^{(1)} = \sum_{k=1}^{\infty} \epsilon_{\left\{\frac{(k-1)M+r}{n},\, a_n^{-1}\left(X_{(k-1)M+r}^{(0)}, ..., X_{(k-1)M+r}^{(M-1)}\right)\right\}}$$

and

$$N_n^{(2)} = \sum_{i=0}^{M-1} \sum_{k=1}^{\infty} \epsilon_{\left\{\frac{(k-1)M+r}{n},\, a_n^{-1} X_{(k-1)M+r}^{(i)} \mathbf{v}_i\right\}} ,$$

where $\mathbf{v}_s$ is the unit vector in $\mathbb{R}^M$ with 1 in the s-th component and the rest zero, differ negligibly, as $n \to \infty$. In doing so we must prove that

$$d\left(N_n^{(1)}, N_n^{(2)}\right) \to 0 , \tag{3.1}$$

in probability, where d is the vague metric on the space of point measures in which $N_n^{(1)}$ and $N_n^{(2)}$ live. Here $N_n^{(2)}$ concentrates all its points on the axes $\mathbf{v}_s$, and (3.1) is expressing the fact that, for each k, at most one of the M components $X_k^{(i)}$

is non-negligible as compared to a_n. From the definition of vague convergence, (3.1) follows if

$$N_n^{(1)}(f) - N_n^{(2)}(f) \to 0 \ , \tag{3.2}$$

in probability for each $f \in C_K^+(E_M)$, the space of continuous non-negative functions with compact support on E_M. To prove (3.2), suppose that f is such a function. Because of compactness, the support of f is contained in the set

$$[0,\xi] \times \Big\{\mathbf{x}\colon\ \mathbf{x} \in [0,\infty]^M \backslash \{\mathbf{0}\},\ \max_{0\le i\le M-1} x_i > \delta\Big\} \ ,$$

for some $\xi > 0$ and $\delta > 0$. Note therefore that f vanishes in $[0,\xi] \times [0,\delta]^M$. For an arbitrary $y \in (0,\delta)$ we define S_y as

$$S_y \Big\{\mathbf{x}\colon\ \mathbf{x} \in [0,\infty]^M \backslash \{\mathbf{0}\},\ \text{at most one component } x_i > y\Big\} \ ,$$

and

$$N_n^{(h)}(f) = \int_{[0,\xi]\times S_y} f\, dN_n^{(h)} + \int_{[0,\xi]\times S_y^c} f\, dN_n^{(h)} \ , \qquad h = 1, 2 \ .$$

Note that

$$E\bigg(\int_{[0,\xi]\times S_y^c} f\, dN_n^{(1)}\bigg) \le$$

$$\le\ (\sup f)\, E\Big(N_n^{(1)}\big([0,\xi] \times S_y^c\big)\Big)$$

$$\le\ (\sup f) \left[\frac{n}{M}\right] \xi\, P\Big[2 \text{ or more } X^{(0)}_{M(k-1)+r}, ..., X^{(M-1)}_{M(k-1)+r} > a_n\, y\Big]$$

$$\le\ (\sup f) \left[\frac{n}{M}\right] \xi \binom{M}{2} P\Big(X^{(i_1)}_{M(k-1)+r} > a_n\, y,\ X^{(i_2)}_{M(k-1)+r} > a_n\, y\Big) \to 0, \quad n \to \infty \ ,$$

which follows by (2.14). Furthermore, it is also true that

$$\int_{[0,\xi]\times S_y^c} f\, dN_n^{(2)} = 0 \ .$$

Thus, in proving (3.2) it is enough to show that

$$\int_{[0,\xi]\times S_y} f\, dN_n^{(1)} - \int_{[0,\xi]\times S_y} f\, dN_n^{(2)} \to 0 \ ,$$

in probability. This last statement follows by the same arguments used in the proof of Proposition 4.26 in Resnick [33]. We skip the details. Consider now the map $T\colon (M_p(E_1))^M \to M_p(E_M)$ such that, for a fixed value of $r = 1..., M$

$$T\left(\sum_{k=1}^{\infty}\sum_{j=1}^{\infty} \epsilon_{\big\{T^{(i)}_{k,r},\, J^{(i)}_{k,r} U_{k,1,r} \cdots U_{k,M(j-1)+i,r}\big\}},\ \ i = 0, ..., M-1\right) =$$

$$= \sum_{i=0}^{M-1}\sum_{k=1}^{\infty}\sum_{j=1}^{\infty} \epsilon_{\big\{T^{(i)}_{k,r},\, J^{(i)}_{k,r} U_{k,1,r} \cdots U_{k,M(j-1)+i,r}\, \mathbf{v}_i\big\}} \ .$$

Note that this map is continuous and hence by the continuous mapping theorem

$$
\begin{aligned}
& T\left(\sum_{k=1}^{\infty} \epsilon_{\left\{\frac{(k-1)M+r}{n},\, a_n^{-1} X^{(i)}_{(k-1)M+r}\right\}},\quad i=0,...,M-1\right) = \\
& \qquad = \sum_{i=0}^{M-1}\sum_{k=1}^{\infty} \epsilon_{\left\{\frac{(k-1)M+r}{n},\, a_n^{-1} X^{(i)}_{(k-1)M+r}\mathbf{v}_i\right\}} \implies \\
(3.3) \qquad & \implies T\left(\sum_{k=1}^{\infty}\sum_{j=1}^{\infty} \epsilon_{\left\{T^{(i)}_{k,r},\, J^{(i)}_{k,r} U_{k,1,r}\cdots U_{k,M(j-1)+i,r}\right\}},\quad i=0,...,M-1\right) = \\
& \qquad = \sum_{i=0}^{M-1}\sum_{k=1}^{\infty}\sum_{j=1}^{\infty} \epsilon_{\left\{T^{(i)}_{k,r},\, J^{(i)}_{k,r} U_{k,1,r}\cdots U_{k,M(j-1)+i,r}\mathbf{v}_i\right\}} ,
\end{aligned}
$$

in $M_p(E_M)$. Finally the map $T\colon M_p(E_M) \to M_p(E_1)$ defined by

$$
T\left(\sum_{k=1}^{\infty} \epsilon_{\left\{\frac{(k-1)M+r}{n},\, a_n^{-1}\left(X^{(0)}_{(k-1)M+r},...,X^{(M-1)}_{(k-1)M+r}\right)\right\}}\right) = \sum_{k=1}^{\infty} \epsilon_{\left\{\frac{(k-1)M+r}{n},\, a_n^{-1} X_{(k-1)M+r}\right\}} ,
$$

is almost surely continuous with respect to the distribution of (3.3). Hence applying the continuous mapping theorem we obtain

$$
\begin{aligned}
& T\left(\sum_{k=1}^{\infty} \epsilon_{\left\{\frac{(k-1)M+r}{n},\, a_n^{-1}\left(X^{(0)}_{(k-1)M+r},...,X^{(M-1)}_{(k-1)M+r}\right)\right\}}\right) = \\
& \qquad = \sum_{k=1}^{\infty} \epsilon_{\left\{\frac{(k-1)M+r}{n},\, a_n^{-1} X_{(k-1)M+r}\right\}} \implies \\
& \implies T\left(\sum_{i=0}^{M-1}\sum_{k=1}^{\infty}\sum_{j=1}^{\infty} \epsilon_{\left\{T^{(i)}_{k,r},\, J^{(i)}_{k,r} U_{k,1,r}\cdots U_{k,M(j-1)+i,r}\mathbf{v}_i\right\}}\right) = \\
& \qquad = \sum_{i=0}^{M-1}\sum_{k=1}^{\infty}\sum_{j=1}^{\infty} \epsilon_{\left\{T^{(i)}_{k,r},\, J^{(i)}_{k,r} U_{k,1,r}\cdots U_{k,M(j-1)+i,r}\right\}} ,
\end{aligned}
$$

providing

$$
\sum_{k=1}^{\infty} \epsilon_{\left\{\frac{k}{n},\, a_n^{-1}X_k\right\}} \implies \sum_{r=1}^{M}\sum_{i=0}^{M-1}\sum_{k=1}^{\infty}\sum_{j=1}^{\infty} \epsilon_{\left\{T^{(i)}_{k,r},\, J^{(i)}_{k,r} U_{k,1,r}\cdots U_{k,M(j-1)+i,r}\right\}} . \qquad \square
$$

The distribution of $M_n = \max_{1\le k\le n}(X_k)$ and its corresponding extremal index can now be obtained.

Corollary 3.1. *Under the conditions of the above theorem,*

1. *as* $n \to \infty$

$$P\big(M_n \le a_n x\big) \;\to\; \exp\Big\{-\frac{1}{M}\, EW^\alpha x^{-\alpha}\Big\} \;, \tag{3.4}$$

with

$$W \;=\; \bigvee_{r=1}^{M} \gamma_{r,1} \bigvee_{i=0}^{M-1} \left(\gamma_{r+M-i-1,r} \bigvee_{j=1}^{\infty} \big\{U_{1,1,r} \cdots U_{1,M(j-1)+i,r}\big\}\right) \;;$$

2. *the periodic sequence* $(X_k)_{k\in\mathbb{Z}}$ *has extremal index*

$$\theta \;=\; \frac{\Big\{1-\prod_{s=1}^{M} EA^\alpha_{r-s}\Big\}\, EW^\alpha}{\sum_{r=1}^{M}\gamma_{r,1}\,\sum_{i=0}^{M-1}\Big(\prod_{s=1}^{i} EA^\alpha_{r-s}\Big)\,\gamma_{r+M-i-1,r}} \;.$$

Proof:

$$P\big(M_n \le a_n x\big) \;=\; P\left(\sum_{k=1}^{\infty} \epsilon_{\left\{\frac{k}{n},\, a_n^{-1}X_k\right\}}\big((0,1]\times(x,\infty]\big) \,=\, 0\right) \;\Longrightarrow$$

$$\Longrightarrow\; P\left(\sum_{r=1}^{M}\sum_{i=0}^{M-1}\sum_{k=1}^{\infty}\sum_{j=1}^{\infty} \epsilon_{\left\{T^{(i)}_{k,r},\, J^{(i)}_{k,r} U_{k,1,r}\cdots U_{k,M(j-1)+i,r}\right\}}\big((0,1]\times(x,\infty]\big) \,=\, 0\right) \;.$$

The event

$$\left\{\sum_{r=1}^{M}\sum_{i=0}^{M-1}\sum_{k=1}^{\infty}\sum_{j=1}^{\infty} \epsilon_{\left\{J^{(i)}_{k,r} U_{k,1,r}\cdots U_{k,M(j-1)+i,r}\right\}}(x,\infty] \,=\, 0\right\} \;,$$

is equivalent to the event that none of the points

$$\Big\{J^{(i)}_{k,r} U_{k,1,r} \cdots U_{k,M(j-1)+i,r}\,,\;\; r=1,...,M,\;\; i=0,...,M-1,\;\; k,j\in\mathbb{N}\Big\} \;,$$

exceeds x. The latter can be expressed as the set

$$\bigcap_{r=1}^{M}\bigcap_{i=0}^{M-1}\bigcap_{k=1}^{\infty} \Big\{J^{(i)}_{k,r} V^{(i)}_{k,r} \le x\Big\} \;, \tag{3.5}$$

where

$$V^{(i)}_{k,r} \;=\; \bigvee_{j=1}^{\infty} \big\{U_{k,1,r}\cdots U_{k,M(j-1)+i,r}\big\} \;.$$

For a fixed value of $r=1...,M$ and $i=0,...,M-1$ it follows that $\{J^{(i)}_{k,r}V^{(i)}_{k,r}\}_{k\in\mathbb{N}}$ are the points of a PRM on $(0,\infty]$ with mean measure

$$(1/M)\;\gamma_{r,1}\;\gamma_{r+M-i-1,r}\;E\big(V^{(i)}_{1,r}\big)^\alpha\, x^{-\alpha} \;,$$

(cf. Resnick, [32]). Since the set in (3.5) can be expressed as

$$\Big\{J^{(i)}_{k,r} W \le x,\;\; r=1,...,M,\; i=0,...,M-1,\; k\in\mathbb{N}\Big\}\;,$$

with

$$W \;=\; \bigvee_{r=1}^{M}\,\bigvee_{i=0}^{M-1} V^{(i)}_{1,r}\;,$$

the set $\big\{J^{(i)}_{k,r} W,\;\; r=1,...,M,\; i=0,...,M-1,\; k\in\mathbb{N}\big\}$ contains the points of a PRM on $(0,\infty)$ with mean measure $EW^\alpha x^{-\alpha}$ and the result follows.

Finally, we concentrate on the examination of the extremal index. Define $(\hat{X}_k)_{k\in\mathbb{Z}}$ as the *associated independent* M-periodic sequence of $(X_k)_{k\in\mathbb{Z}}$, i.e., $\hat{X}_1, \hat{X}_2...$, are independent random variables being the tail distribution of $\hat{X}_r$ as in (2.13), for $r=1,...,M$. Further we define $M_n^{\hat{X}} = \max_k(\hat{X}_k)$. From Theorem 2.1 and classical extreme value theory we obtain that

$$\text{(3.6)}\qquad \begin{aligned} &P\big(M_n^{\hat{X}} \le a_n x\big) \;\to \\ &\to\; \exp\left\{-\frac{1}{M}\sum_{r=1}^{M}\tau_r\left(\frac{1}{1-\prod_{s=1}^{M} EA^\alpha_{r-s}}\sum_{i=0}^{M-1}\left(\prod_{s=1}^{i} EA^\alpha_{r-s}\right)\gamma_{r+M-i-1,r}\right)\right\}\,. \end{aligned}$$

By comparing (3.4) with (3.6) the expression of the extremal index is obtained; see Leadbetter *et al.* [26]. □

4. EXAMPLES

Consider that X_k is given in the form

$$X_k \;=\; \sum_{j=1}^{\infty}\left(\prod_{s=1}^{j-1} b\,Z_{k-s}\right) b\,Z^2_{k-j}\;, \qquad k\in\mathbb{Z}\;,$$

with $b>0$ a positive constant. Note that the process $(Y_k)_{k\in\mathbb{Z}}$ defined as

$$Y_k \;=\; X_k + Z_k\;,$$

satisfies the bilinear recursion

$$Y_k \;=\; b\,Y_{k-1}\,Z_{k-1} + Z_k\;, \qquad k\in\mathbb{Z}\;.$$

The reason in considering the tail behavior of X_k rather that Y_k itself is due to the fact that the contribution of the term Z_k on the extremal behavior of Y_k is negligible.

For deriving probabilistic and extremal properties of this process we will make extensive use of the fact that X_k can be embedded in the form (2.4) if $(A_k, B_k) = (bZ_k, bZ_k^2)$. We further assume that

$$\bar{F}_r(x) = P\big(Z_r^2 > x\big) = x^{-\alpha/2} L_r(x) , \qquad r = 1, ..., M ,$$

and that

$$b^{\alpha/2} E Z_r^{\alpha/2} < 1 , \qquad r = 1, ..., M .$$

It follows by Lemma 2.2 and the fact that

$$P\big(B_r > x\big) = P\big(b Z_r^2 > x\big) = b^{\alpha/2} x^{-\alpha/2} L_r(x) ,$$

for $r = 1, ..., M$ and $i = 0, ..., M-1$

$$\lim_{x\to\infty} \frac{P\big(X_r^{(i)} > x\big)}{P\big(Z_r^2 > x\big)} = \frac{b^{(i+1)\alpha/2} \prod_{s=1}^{i} E Z_{r-s}^{\alpha/2}}{1 - b^{M\alpha/2} \prod_{s=1}^{M} E Z_{r-s}^{\alpha/2}} \gamma_{r+M-i-1,r} .$$

Furthermore, by Theorem 2.1, for $r = 1, ..., M$

$$\lim_{x\to\infty} \frac{P\big(X_r > x\big)}{P\big(Z_r^2 > x\big)} = \frac{b^{\alpha/2}}{1 - b^{M\alpha/2} \prod_{s=1}^{M} E Z_{r-s}^{\alpha/2}} \sum_{i=0}^{M-1} \left(\prod_{s=1}^{i} E Z_{r-s}^{\alpha/2}\right) b^{\alpha i/2} \gamma_{r+M-i-1,r} .$$

The expression of the extremal index can be calculated from Corollary 3.1, providing

$$\theta = \frac{\Big(1 - b^{M\alpha/2} \prod_{s=1}^{M} E Z_{r-s}^{\alpha/2}\Big) E W^{\alpha/2}}{b^{\alpha/2} \sum_{r=1}^{M} \gamma_{r,1} \sum_{i=0}^{M-1} \Big(\prod_{s=1}^{i} E Z_{r-s}^{\alpha/2}\Big) \gamma_{r+M-i-1,r}\, b^{\alpha i/2}} .$$

Extensions for bivariate bilinear models can be easily obtained from the previous results; see Kumar [25] for details.

ACKNOWLEDGMENTS

Research (partially) supported by Unidade de Investigação Matemática e Aplicações of Universidade de Aveiro through Programa Operacional "Ciência, Tecnologia, Inovação" of the Fundação para a Ciência e a Tecnologia (FCT) cofinanced by the European Community fund FEDER. The author would also like to express his gratitude to the referee for all helpful comments.

REFERENCES

[1] ALPUIM, M.T. (1988). *Contribuições à Teoria de Extremos em Sucessões Dependentes*, Unpublished doctoral dissertation, University of Lisbon.

[2] BABILLOT, M.; BOUGEROL, P. and ELIE, L. (1997). The random difference equation $X_n = A_n X_{n-1} + B_n$ in the critical case, *Ann. Probab.*, **25**, 478–493.

[3] BALLERINI, R. (1995). Nonconstant levels for T-year return periods for two-dimensional Poisson processes with periodic intensity, *Statist. Probab. Lett.*, **22**, 43–47.

[4] BALLERINI, R. and MCCORMICK, W.P. (1989). Extreme value theory for processes with periodic variances, *Stochastic Models*, **5**, 1403–1433.

[5] BALLERINI, R. and WAYLEN, P. (1989). Extreme precipitation events generated by periodic processes, *Water Resour. Res.*, **25**, 1403–1411.

[6] BAXENDALE, P.H. and KHASMINSKII, R.Z. (1998). Stability index for products of random transformations, *Adv. Appl. Probab.*, **30**, 968–988.

[7] BOUGEROL, P. and PICARD, N. (1992). Strict stationarity of generalized autoregressive processes, *Ann. Probab.*, **20**, 1714–1730.

[8] BRANDT, A.; FRANKEN, P. and LISEK, B. (1990). *Stationary Stochastic Models*, Wiley, Chichester.

[9] COLES, S.G. (2001). *An Introduction to Statistical Modeling of Extreme Values*, Springer-Verlag, London.

[10] DAVIS, R.A. and RESNICK, S.I. (1996). Limit theory for bilinear processes with heavy-tailed noise, *Ann. Appl. Probab.*, **6**, 1191–1210.

[11] DE HAAN, L.; RESNICK, S.I.; ROOTZÉN, H. and DE VRIES, C.G. (1989). Extremal behavior of solutions to a stochastic difference equation with applications to ARCH processes, *Stoch. Process. Appl.*, **32**, 213–224.

[12] DUFRESNE, D. (1991). The distribution of a perpetuity, with applications to risk theory and pension funding, *Scand. Actuar. J.*, 39–79.

[13] EMBRECHTS, P.; KLÜPPELBERG, C. and MIKOSCH, T. (1997). *Modelling Extremal Events for Insurance and Finance*, Springer-Verlag, Heildelberg.

[14] FEIGIN, P.; KRATZ, M. and RESNICK, S.I. (1996). Parameter estimation for moving averages with positive innovations, *Ann. Appl. Probab.*, **6**, 1157–1190.

[15] FERREIRA, H. (1994). Multivariate extreme values in T-periodic random sequences under mild oscilation restrictions, *Stoch. Process. Appl.*, **49**, 111–125.

[16] FERREIRA, H. (1995). Extremes of a random number of variables from periodic sequences, *J. Stat. Plann. Inf.*, **45**, 133–141.

[17] FERREIRA, H. and MARTINS, A.P. (2003). The extremal index of sub-sampled periodic sequences with strong local dependence, *Revstat – Statistical Journal*, **1**, 15–24.

[18] GOLDIE, C.M. (1991). Implicit renewal theory and tails of solutions of random equations, *Ann. Appl. Probab.*, **1**, 126–166.

[19] GREY, D.R. (1994). Regular variation in the tail behavior of solutions of random difference equations, *Ann. Appl. Probab.*, **4**, 169–183.

[20] Grincevičius, A. K. (1975). One limit distribution for a random walk on the line, *Lithuanian Math. J.*, **15**, 580–589.

[21] Horowitz, J. (1980). Extreme values for a nonstationary processes: an application to air quality analysis, *Technometrics*, **22**, 469–478.

[22] Kallenberg, O. (1983). *Random Measures*, Akademie-Verlag, Berlin.

[23] Kesten, H. (1973). Random difference equations and renewal theory for products of random matrices, *Acta Math.*, **131**, 207–248.

[24] Konstantinides, D.G. and Mikosch, T. (2005). Large deviations and ruin probabilities for solutions to stochastic recurrence equations with heavy-tailed innovations, *Ann. Probab.*, **33**, 1992–2035.

[25] Kumar, K. (1988). *Bivariate bilinear models and their specification.* In "Non-Linear Time Series and Signal Processing" (R. Mohler, Ed.), Springer-Verlag, U.S.A., 59–74.

[26] Leadbetter, M.R.; Lindgren, G. and Rootzén, H. (1983). *Extremes and Related Properties of Random Sequences and Processes*, Springer-Verlag, New York.

[27] Martins, A.P. and Ferreira, H. (2004). Extremes of periodic moving averages of random variables with varying tail probabilities, *Sort*, **28**, 161–176.

[28] Mikosch, T. (2003). *Modeling dependence and tails of financial time series.* In "Extreme Values in Finance, Telecommunications, and the Environment" (B. Finkenstaedt and H. Rootzén, Eds.), Chapman and Hall, 185–286.

[29] Niu, X.F. (1996). Nonlinear additive models for environmental time series, with applications to ground-level ozone data analysis, *J. Amer. Statist. Soc.*, **91**, 1310–1321.

[30] Niu, X.F. (1997). Extreme value for a class of nonstationary time series with applications, *Ann. Appl. Probab.*, **7**, 508–522.

[31] Perfekt, R. (1994). Extremal behaviour of stationary Markov chains with applications, *Ann. Appl. Probab.*, **4**, 529–548.

[32] Resnick, S.I. (1986). Point process, regular variation and weak convergence, *Adv. Appl. Probab.*, **18**, 66–138.

[33] Resnick, S.I. (1987). *Extreme Values, Point Processes and Regular Variation*, Springer-Verlag, New-York.

[34] Resnick, S.I. and Van den Berg, E. (2000). Sample correlation behavior for the heavy tailed general bilinear process, *Commun. Statist. – Stochastic Models*, **16**, 233–258.

[35] Scotto, M.G. (2005). Extremes of a class of deterministic sub-sampled processes with applications to stochastic difference equations, *Stoch. Proc. Appl.*, **115**, 417–434.

[36] Stărică, C. (1999). Multivariate extremes for models with constant conditional correlations, *J. Empirical Finance*, **6**, 515–553.

[37] Vervaat, W. (1979). On the stochastic difference equation and a representation on non-negative infinitely random variables, *Adv. Appl. Probab.*, **11**, 750–783.

REVSTAT – Statistical Journal
Volume 5, Number 3, November 2007, 249–267

LARGE DEVIATIONS AND BERRY–ESSEEN INEQUALITIES FOR ESTIMATORS IN NONLINEAR NONHOMOGENEOUS DIFFUSIONS

Authors: Jaya P.N. Bishwal
– Department of Mathematics and Statistics,
University of North Carolina at Charlotte,
376 Fretwell Bldg., 9201 University City Blvd.,
Charlotte, NC 28223, U.S.A.
J.Bishwal@uncc.edu

Received: February 2007 Revised: June 2007 Accepted: September 2007

Abstract:

- Bounds on the large deviations probability of the maximum likelihood estimator and regular Bayes estimators, and Berry–Esseen type bound for the suitably normalized maximum likelihood estimator of a parameter appearing nonlinearly in the nonhomogeneous drift coefficient of Itô stochastic differential equation are obtained under some regularity conditions. Berry–Esseen results are illustrated for nonhomogeneous Ornstein–Uhlebeck process.

Key-Words:

- *nonhomogeneous diffusion processes; Itô stochastic differential equation; drift parameter; maximum likelihood estimator; Bayes estimators; large deviations probability; Berry–Esseen type inequality.*

AMS Subject Classification:

- 62M05, 62F12, 60H10, 60J60.

1. INTRODUCTION

Nonhomogenous diffusions are useful for modeling term structure of interest rates in finance and other fields. Asymptotic properties such as weak consistency, asymptotic normality and convergence of moments of maximum likelihood estimator (MLE) and Bayes estimators (BEs) of the drift parameter in the nonlinear nonhomogeneous Itô stochastic differential equations having nonstationary solutions were first studied by Kutoyants (1978) for the small noise asymptotics case and Kutoyants (1984) for the general case which includes both small noise and long time asymptotics. The approach was through Ibragimov and Khasminskii (1981). Later on, strong consistency and asymptotic normality for large sample case were studied by Borkar and Bagchi (1982), Mishra and Prakasa Rao (1985a) and Levanony, Shwartz and Zeitouni (1994) using the martingale approach under stronger regularity conditions. Asymptotic normality of BEs was studied by Mishra (1989) and Harison (1992) as a consequence of Bernstein-von Mises theorem. Slightly weaker assumptions than those used in Kutoyants (1984) were used by Yoshida (1990) to obtain the asymptotic behaviour of M-estimator. For first order theory in general nonergodic stochastic models through the LAMN (defined below) approach, see Basawa and Scott (1983). See the monograph Bishwal (2007) for recent results on likelihood asymptotics and Bayesian asymptotics for drift estimation of finite and infinite dimensional stochastic differential equations.

All the above results are on first order asymptotics. Beyond the first order asymptotics in consistency, Florens and Pham (1999) obtained large deviations for MLE and a minimum contrast estimator for the Ornstein–Uhlenbeck process. For the nonlinear stationary homogeneous diffusions a large deviations upper bound for the MLE and Bayes estimators was obtained by Bishwal (1999). For the nonhomogeneous diffusions, Levanony (1994) obtained the conditional large deviations upper and lower bounds for the MLE through the martingale approach following Dupuis and Kushner (1989). He obtained unconditional large deviations lower bounds following Bahadur *et al.* (1980). We obtain unconditional large deviations upper bounds following Ibragimov and Khasminskii (1981).

Beyond the first order results in asymptotic normality, Berry–Esseen type bounds in the linear homogeneous case were obtained by Mishra and Prakasa Rao (1985b) which were sharpened to the Ornstein–Uhlenbeck process by Bose (1986), Bishwal and Bose (1995) and Bishwal (2000a) respectively in order. Sharp Berry–Esseen bound for the Bayes estimators and minimum contrast estimator were obtained in Bishwal (2000b) and Bishwal (2005) respectively. In the above works on Ornstein–Uhlenbeck process, stationarity was not assumed. For nonlinear stationary ergodic diffusion, Edgeworth expansion of the distribution of the MLE was obtained by Yoshida (1997) and that for M-estimator by Sakamoto and Yoshida (1998) through the Malliavin calculus approach. As far as we know, no result is known on the rate of convergence to normality of the MLE in the nonergodic case. We obtain a Berry–Esseen type bound for the MLE following Michel and Pfanzagl (1971). Finally Berry–Esseen results are illustrated for a nonhomogeneous Ornstein–Uhlenbeck process.

2. MODEL, ASSUMPTIONS AND PRELIMINARIES

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0} P)$ be a stochastic basis satisfying the usual hypotheses on which is defined a diffusion process $\{X_t,\, t \geq 0\}$ satisfying the Itô stochastic differential equation

$$dX_t = f(\theta, t, X_t)\, dt + dW_t\ , \qquad t \geq 0, \quad X_0 = 0 \tag{2.1}$$

where $\{W_t,\, t \geq 0\}$ is a standard Wiener process, $f(\theta, t, x)$ is a known real valued function continuous on $\Theta \times [0,T] \times \mathbb{R}$ where Θ is a closed interval of the real line and the parameter θ is unknown, which is to be estimated on the basis of observation of the process $\{X_t,\, 0 \leq t \leq T\} =: X_0^T$. Let θ_0 be the true value of the parameter which lies inside the parameter space Θ.

Let P_θ^T be the measure generated by the process X_0^T on the space (C_T, B_T) of continuous functions on $[0,T]$ with the associated Borel σ-algebra B_T associated to the sup-norm topology of C_T. Let E_θ^T be the expectation with respect to the measure P_θ^T. Suppose P_θ^T is absolutely continuous with respect to $P_{\theta_0}^T$. Then it is well known that (see Liptser and Shiryayev (1977, p. 239)

$$\begin{aligned} L_T(\theta) &:= \frac{dP_\theta^T}{dP_{\theta_0}^T}(X_0^T) \\ &= \exp\left\{ \int_0^T \Big[f(\theta, s, X_s) - f(\theta_0, s, X_s) \Big]\, dW_s \right. \\ &\quad \left. - \frac{1}{2} \int_0^T \Big[f(\theta, s, X_s) - f(\theta_0, s, X_s) \Big]^2 ds \right\} . \end{aligned} \tag{2.2}$$

is the Radon–Nikodym derivative (likelihood) of P_θ^T with respect to $P_{\theta_0}^T$. The MLE θ_T of θ based on X_0^T is defined as

$$\theta_T := \arg\max_{\theta\in\Theta} L_T(\theta)\ .$$

Throughout the paper prime denotes derivative with respect to θ. Let us denote the log-likelihood function by $l_t(\theta) \equiv \log L_T(\theta)$, and let $l_t'(\theta) \equiv U_T(\theta)$, $l_T''(\theta) \equiv H_T(\theta)$ and $l_T'''(\theta) \equiv Q_T(\theta)$.

If $L_T(\theta)$ is continuous in θ, it can be shown that there exists a measurable MLE by using Lemma 3.3 in Schmetterer (1974). Hereafter, we assume the existence of such a measurable MLE. We assume the following regularity conditions on $f(\theta, t, x)$.

(**A1**) $P_{\theta_1} \neq P_{\theta_2}$ for $\theta_1 \neq \theta_2$ in Θ.

(**A2**) $\{X_t\}$ is the unique strong solution of (2.1) with

$$P_\theta\left(\int_0^T f^2(\theta,t,X_t)\,dt < \infty\right) = 1 \qquad \text{for all } \theta\in\Theta,\ \ T<\infty\ . \tag{2.3}$$

The condition (2.3) ensures that $P_\theta^T \ll P_W^T$ for all θ where P_W^T is the standard Wiener measure and likelihood function is given by

$$\frac{dP_\theta^T}{dP_W^T} = \exp\left\{\int_0^T f(\theta,t,X_t)\,dX_t - \frac{1}{2}\int_0^T f^2(\theta,t,X_t)\,dt\right\}. \tag{2.4}$$

(**A3**) (**i**) $f(\theta,t,x)$ is differentiable in t and x.
The log-likelihood with respect to P_W^T can be written as

$$\log\frac{dP_\theta^T}{dP_W^T} = \int_0^T f(\theta,t,X_t)\,dX_t - \frac{1}{2}\int_0^T f^2(\theta,t,X_t)\,dt\ . \tag{2.5}$$

(**ii**) The integrals in (2.4) and (2.5) can be differentiated twice under the integral sign with respect to θ.
Let $I_T(\theta) := \int_0^T f'^2(\theta,t,X_t)\,dt$ and $Y_T(\theta) := \int_0^T f''^2(\theta,t,X_t)\,dt$.

(**iii**) l_T'' is continuous in a neighborhood V_θ of θ for every $\theta\in\Theta$ and

$$n_T = n_T(\theta) := E_\theta(I_T(\theta)) < \infty\ , \qquad E_\theta(Y_T(\theta)) < \infty$$

with $n_T\to\infty$ as $T\to\infty$ and there exists a constant C_0 such that for any $\theta,\theta_1,\theta_2\in\Theta$

$$\frac{E_\theta(I_T(\theta_2))}{n_T(\theta_1)} < C_0\ .$$

(**iv**) $\dfrac{I_T(\theta)}{n_T} \xrightarrow{P_\theta} 1$ as $T\to\infty$.

(**A4**) Suppose there exists $\gamma\geq 2$ and $C>0$ such that for all $\theta\in\Theta$

$$E_\theta\exp\left\{-\frac{1}{3}\int_0^T\Big[f\big(\theta+un_T^{-1/2},t,X_t\big)-f(\theta,t,X_t)\Big]^2 dt\right\} \leq \exp\big(-C|u|^\gamma\big)\,.$$

(**A5**) Suppose that there exists $m_T = m_T(\theta)\uparrow\infty$ as $T\to\infty$ such that

(**i**) $\dfrac{I_T(\theta)}{m_T}\xrightarrow{P_\theta}\eta(\theta)$ as $T\to\infty$ where $P_\theta\big(\eta(\theta)>0\big)>0$ and $E(\eta^{-1}(\theta))<\infty$.

(**ii**) $\dfrac{Y_T(\theta)}{m_T}\xrightarrow{P_\theta}\xi(\theta)$ as $T\to\infty$.

Some of regularity conditions (A1)–(A5) can be found in the literature, for example, Borkar and Bagchi (1982) and Levanony *et al.* (1994). However both proved strong consistency and Levanony *et al.* (1994) proved asymptotic normality. We need stronger regularity condition (A4) in order to prove large deviations.

Let us introduce the Bayes estimator. Let Λ be a prior probability measure on $(\Theta, \mathcal{B})$ where $\mathcal{B}$ is the σ-algebra of Borel subsets of Θ. Suppose that Λ has a density $\lambda(\cdot)$ with respect to the Lebesgue measure on $\mathbb{R}$, which is continuous and positive on Θ and possesses a polynomial majorant in Θ.

Let $p(\theta|X_0^T)$ be the posterior density of θ given X_0^T. By Bayes theorem $p(\theta|X_0^T)$ is given by

$$p(\theta|X_0^T) = \frac{L_T(\theta)\,\lambda(\theta)}{\int_\Theta L_T(\theta)\,\lambda(\theta)\,d\theta} .$$

Let $l(\cdot,\cdot)\colon \Theta\times\Theta \to \mathbb{R}$ be a loss function as defined in Ibragimov and Khasminskii (1981) which satisfies the following conditions:

(**B1**) $\psi(u,v) = \psi(u-v)$.

(**B2**) $\psi(u)$ is defined and nonnegative on $\mathbb{R}, \psi(0) = 0$ and $\psi(u)$ is continuous at $u = 0$ but is not identically equal to 0.

(**B3**) ψ is symmetric, i.e., $\psi(u) = \psi(-u)$.

(**B4**) $\{u\colon \psi(u) < c\}$ are convex sets and are bounded for all $c > 0$ sufficiently small.

(**B5**) There exists numbers $\gamma > 0$, $h_0 \geq 0$ such that for $h \geq h_0$

$$\sup\Big\{\psi(u)\colon\ |u| \leq h^\gamma\Big\} \leq \inf\Big\{\psi(u)\colon\ |u| \geq h\Big\} .$$

Clearly, all power loss functions of the form $|u-v|^r$, $r > 0$, satisfy the condition (B1)–(B5). In particular, quadratic loss function $|u-v|^2$ satisfies these conditions.

A Bayes estimator $\widetilde{\theta}_T$ of θ with respect to the loss function $\psi(\theta,\phi)$ and prior density $\lambda(\theta)$ is one which minimizes the posterior risk and is given by

$$\widetilde{\theta}_T := \arg\min_{\phi\in\Theta} \int_\Theta l(\phi,\theta)\, p(\theta|X_0^T)\,d\theta .$$

In particular, for the quadratic loss function $\psi(u,v) = |u-v|^2$, the Bayes estimator $\widetilde{\theta}_T$ becomes the posterior mean given by

$$\widetilde{\theta}_T = \frac{\int_\Theta \phi\ p(\phi|X_0^T)\,du}{\int_\Theta p(\phi|X_0^T)\,d\phi} .$$

Let us consider the likelihood ratio process

$$Z_T(u) := \frac{dP_{\theta+un_T^{-1/2}}}{dP_\theta}(X_0^T) .$$

By (2.2) with $g_t(u) := f\big(\theta + u n_T^{-1/2}, t, X_t\big) - f(\theta, t, X_t)$, we have

$$\begin{aligned} Z_T(u) &= \exp\left\{\int_0^T \Big[f\big(\theta + u n_T^{-1/2}, t, X_t\big) - f(\theta, t, X_t)\Big]\, dW_t \right. \\ &\quad \left. - \frac{1}{2}\int_0^T \Big[f\big(\theta + u n_T^{-1/2}, t, X_t\big) - f(\theta, t, X_t)\Big]^2 dt\right\} \\ &= \exp\left\{\int_0^T g_t(u)\, dW_t - \frac{1}{2}\int_0^T g_t^2(u)\, dt\right\} . \end{aligned}$$

We define the LAMN condition below:

Definition (Le Cam and Yang (1990), Jeganathan (1982, 1995))**.** Let $\mathcal{E}_n = \big(\Omega_n, \mathcal{F}_n, (P_\theta^n, \theta \in \Theta)\big)$, $n \geq 1$, be a sequence of statistical experiments, where Θ is an open subset of $\mathbb{R}$. We denote by

$$\Lambda_{\eta\theta}^n = \log\left(\frac{dP_\eta^n}{dP_\theta^n}\right)$$

the log-likelihood between η and θ at stage n.

We say that the sequence $\mathcal{E}_n$ satisfies the *local asymptotically quadratic* (LAQ) condition at a point $\theta \in \Theta$ if there are random variables Δ_n and Γ_n defined on $(\Omega_n, \mathcal{F}_n)$, $\Gamma_n > 0$ a.s. $[P_\theta^n]$ and a positive numerical sequence $\phi_n \downarrow 0$ such that for each bounded sequence of numbers u_n,

$$\Lambda_{\theta+\phi_n u_n, \theta}^n - \Big(u_n \Delta_n - \frac{1}{2} u_n^2 \Gamma_n\Big) \overset{P_\theta^n}{\to} 0$$

and

$$(\Delta_n, \Gamma_n) \to (\Delta, \Gamma) \qquad \text{in } \ P_\theta^n\text{-distribution}$$

where Δ and Γ are random variables on a measurable space $(\Omega, \mathcal{F}, P)$ with $\Gamma > 0$ a.s. (P) and

$$E_P \exp\Big(u\Delta - \frac{1}{2}u^2\Gamma\Big) = 1 \ .$$

The sequence of experiments is called *locally asymptotically Brownian functional* (LABF) if $\Delta = \int_0^1 F_s\, dW_s$ and $\Gamma = \int_0^1 F_s^2\, ds$ with W a standard Brownian motion and F a predictable process with respect to some filtration in $\mathcal{F}$. It is called *locally asymptotically mixed normal* (LAMN) if $\Delta = \Gamma^{1/2} W_1$ with W_1 standard normal variable independent of Γ and *locally asymptotically normal* (LAN) if, in addition, Γ is nonrandom.

Let $\Phi(.)$ denote the standard normal distribution function and C denote a generic positive constant. We shall use the following lemmas to prove our main results. The first lemma is a revised version of Theorem 19 of Ibragimov and Khasminskii (1981, p. 372) from Kallianpur and Selukar (1993, p. 330).

Lemma 2.1. *Let $\zeta(t)$ be a real valued random function defined on a closed subset F of the Euclidean space $\mathbb{R}^k$. We shall assume that the random process $\zeta(t)$ is measurable and separable. Assume that the following condition is fulfilled: there exist numbers $m \geq r > k$ and a function $H(x)\colon \mathbb{R}^k \to \mathbb{R}^1$ bounded on compact sets such that for all $x, h \in F$, $x+h \in F$,*

$$E\,|\zeta(x)|^m \leq H(x)\ ,$$
$$E\big|\zeta(x+h) - \zeta(x)\big|^m \leq H(x)\,|h|^r\ .$$

Then with probability one the realizations of $\zeta(t)$ are continuous functions of F. Moreover, set

$$w(\delta, \zeta, L) = \sup_{\substack{x,y \in F \\ |x|,|y| \leq L \\ |x-y| \leq \delta}} \big|\zeta(x) - \zeta(y)\big|\ ,$$

then

$$E\big(w(h; \zeta, L)\big) \leq B_0 \Big(\sup_{|x|<L} H(x)\Big)^{1/m} L^k\, h^{(r-k)/m} \log(h^{-1})$$

where $B_0 = 64^k \big(1 - 2^{-(r-k)m}\big)^{-1} + \big(2^{(m-r)/m} - 1\big)^{-1}$.

Lemma 2.2 (Ibragimov and Khasminskii (1981, p. 45))**.** *Let $Z_{\epsilon,\theta}(u)$ be the likelihood ratio function corresponding to the points $\theta + \phi(\epsilon) u$ and θ where $\phi(\epsilon)$ denotes a normalizing factor such that $|\phi(\epsilon)| \to 0$ as $\epsilon \to 0$. Thus $Z_{\epsilon,\theta}$ is defined on the set $U_\epsilon = (\phi(\epsilon))^{-1}(\Theta - \theta)$. Let $Z^\theta_{\epsilon,0}(u)$ possesses the following properties: given a compact set $K \subset \Theta$ there exist numbers $M_1 > 0$ and $m_1 \geq 0$ and functions $g^K_\epsilon(y) = g_\epsilon(y)$ correspond such that*

(1) *For some $\alpha > 0$ and all $\theta \in K$,*

$$\sup_{\substack{|u_1| \leq R \\ |u_2| \leq R}} |u_2 - u_1|^{-\alpha}\, E^{(\epsilon)}_\theta \big|Z^{1/2}_{\epsilon,\theta}(u_2) - Z^{1/2}_{\epsilon,\theta}(u_1)\big|^2 \leq M_1 (1 + R^{m_1})\ .$$

(2) *For all $\theta \in K$ and $u \in U_\epsilon$, $E^{(\epsilon)}_\theta Z^{1/2}_{\epsilon,\theta}(u) \leq e^{-g_\epsilon(u)}$.*

(3) *$g_\epsilon(u)$ is a monotonically increasing to ∞ function of y*

$$\lim_{\substack{y \to \infty \\ \epsilon \to 0}} y^N e^{-g_\epsilon(y)} = 0\ .$$

Let $\{\widetilde{\theta}_\epsilon\}$ be a family of Bayes estimators with respect to the prior density q, which is continuous and positive on K and possesses in Θ a polynomial majorant and a loss function $\omega_\epsilon(u, v) := \psi\big((\phi(\epsilon))^{-1}(u - v)\big)$ where ψ satisfies (B1)–(B5). Then for all N,

$$\lim_{\substack{h \to \infty \\ \epsilon \to 0}} h^N \sup_{\theta \epsilon K} P^{(\epsilon)}_\theta \Big\{\big|(\phi(\epsilon))^{-1}(\widetilde{\theta}_\epsilon - \theta)\big| > h\Big\} = 0\ .$$

If in addition, $\psi(u) = \tau(|u|)$, then for all ϵ sufficiently small, $0 < \epsilon < \epsilon_0$,

$$\sup_{\theta \in K} P_\theta^{(\epsilon)} \Big\{ \big|(\phi(\epsilon))^{-1} (\widetilde{\theta}_\epsilon - \theta)\big| > h \Big\} \le B_0 \, e^{-b_0 g_\epsilon(h)} .$$

Lemma 2.3. *Under the assumptions (A1)–(A4),*

(**a**) $\displaystyle \sup_{\theta \in \Theta} E_\theta^T \Big[Z_T^{1/2}(u_1) - Z_T^{1/2}(u_2) \Big]^2 \le \frac{C_0^2}{4} (u_2 - u_1)^2 \, ;$

(**b**) $\displaystyle \sup_{\theta \in \Theta} E_\theta^T \Big[Z_T^{1/2}(u) \Big] \le C \exp(-C|u|^\gamma) \, .$

Proof: Observe that

$$\begin{aligned} & E_\theta^T \Big[Z_T^{1/2}(u_1) - Z_T^{1/2}(u_2) \Big]^2 = \\ & \qquad = E_\theta^T \big[Z_T(u_1) \big] + E_\theta^T \big[Z_T(u_2) \big] - 2\, E_\theta^T \Big[Z_T^{1/2}(u_1)\, Z_T^{1/2}(u_2) \Big] \\ & \qquad \le 2 - 2\, E_\theta^T \Big[Z_T^{1/2}(u_1)\, Z_T^{1/2}(u_2) \Big] . \end{aligned} \tag{2.6}$$

From Gikhman and Skorohod (1972, p. 82), for all u, we have

$$E_\theta^T \big[Z_T(u) \big] = E_\theta^T \left[\exp\left\{ \int_0^T g_t(u)\, dW_t - \frac{1}{2} \int_0^T g_t^2(u)\, dt \right\} \right] \le 1 . \tag{2.7}$$

Let

$$\begin{aligned} & \theta_1 := \theta + u_1 n_T^{-1/2} , \qquad \theta_2 := \theta + u_2\, n_T^{-1/2} , \\ & \delta_t := f(\theta_2, t, X_t) - f(\theta_1, t, X_t) , \\ & J(\theta_1, \theta_2) := E_{\theta_1}\big(I_T(\theta_2) \big) , \\ & V_T := \exp\left\{ \frac{1}{2} \int_0^T \delta_t \, dW_t - \frac{1}{4} \int_0^T \delta_t^2 \, dt \right\} = \left(\frac{dP_{\theta_2}^T}{dP_{\theta_1}^T} \right)^{1/2} . \end{aligned} \tag{2.8}$$

By Itô formula, V_T can be represented as

$$V_T = 1 - \frac{1}{8} \int_0^T V_t \, \delta_t^2 \, dt + \frac{1}{2} \int_0^T V_t \, \delta_t \, dW_t \, . \tag{2.9}$$

The random process $\big\{ V_t^2, \mathcal{F}_t, P_\theta^T, \, 0 \le t \le T \big\}$ is a martingale and from the

$\mathcal{F}_t$-measurability of δ_t for each $t \in [0,T]$,

$$
\begin{aligned}
E_{\theta_1}^T \int_0^T V_t^2\, \delta_t^2\, dt &= E_{\theta_1}^T \int_0^T E_{\theta_1}^T (V_t^2 \,|\, \mathcal{F}_t)\, \delta_t^2\, dt \\
&= E_{\theta_1}^T V_T^2 \int_0^T \delta_t^2\, dt \\
&= \int V_T^2 \left(\int_0^T \delta_t^2\, dt \right) dP_{\theta_1} \\
&= \int \left(\int_0^T \delta_t^2\, dt \right) dP_{\theta_2}^T \qquad (2.10)\\
&= E_{\theta_2}^T \left(\int_0^T \delta_t^2\, dt \right) \\
&= E_{\theta_2}^T \int_0^T \big| f(\theta_2, t, X_t) - f(\theta_1, t, X_t) \big|^2\, dt \\
&= E_{\theta_2}^T \int_0^T \left(\int_{\theta_1}^{\theta_2} f'(y,t,X_t)\, dy \right)^2 dt \qquad \text{(by (A1))} \\
&\le (\theta_2 - \theta_1)\, E_{\theta_2}^T \int_0^T \int_{\theta_1}^{\theta_2} f'^2(y,t,X_t)\, dy\, dt \qquad (2.11)\\
&= (\theta_2 - \theta_1) \int_{\theta_1}^{\theta_2} J(\theta_2, y)\, dy \; < \; \infty \; .
\end{aligned}
$$

Hence $E_{\theta_1}^T \int_0^T V_t\, \delta_t\, dW_t = 0$. Therefore, using $|a\,b| \le \frac{a^2+b^2}{2}$, we obtain from (2.10)

$$
\begin{aligned}
E_{\theta_1}^T (V_T) &= 1 - \frac{1}{8} \int_0^T E_{\theta_1}^T (\delta_t V_t \,.\, \delta_t)\, dt \\
&\ge 1 - \frac{1}{16} \int_0^T E_{\theta_1}^T \delta_t^2\, dt - \frac{1}{16} \int_0^T E_{\theta_1}^T V_t^2\, \delta_t^2\, dt \qquad (2.12)\\
&= 1 - \frac{1}{16}\, E_{\theta_1}^T \int_0^T \delta_t^2\, dt - \frac{1}{16}\, E_{\theta_2}^T \int_0^T \delta_t^2\, dt \qquad \text{(by (2.11))} \; .
\end{aligned}
$$

Now

$$
\begin{aligned}
E_\theta^T \Big[Z_T^{1/2}(u_1)\, Z_T^{1/2}(u_2) \Big] &= E_\theta^T \left[\frac{dP^T_{\theta + u_1 n_T^{-1/2}}}{dP_\theta^T} \right]^{1/2} \left[\frac{dP^T_{\theta + u_2 n_T^{-1/2}}}{dP_\theta^T} \right]^{1/2} \\
&= \int \left[\frac{dP_{\theta_1}^T}{dP_\theta^T} \right]^{1/2} \left[\frac{dP_{\theta_2}^T}{dP_\theta^T} \right]^{1/2} dP_\theta^T \qquad (2.13)\\
&= \int \left[\frac{dP_{\theta_2}^T}{dP_{\theta_1}^T} \right]^{1/2} dP_{\theta_1}^T \;=\; E_{\theta_1}^T (V_T) \; .
\end{aligned}
$$

Substituting (2.13) into (2.6) and using (2.12), we obtain

$$
\begin{aligned}
E_\theta\Big[Z_T^{1/2}(u_1) - Z_T^{1/2}(u_2)\Big]^2 \le & \\
&\le\ 2 - 2\,E_{\theta_1}(V_T) \\
&\le\ \frac{1}{8}\,E_{\theta_1}\int_0^T \delta_t^2\,dt + \frac{1}{8}\,E_{\theta_2}\int_0^T \delta_t^2\,dt \\
&\le\ \frac{1}{8}\,(\theta_2-\theta_1)\int_{\theta_1}^{\theta_2}\Big[J(\theta_1,y)+J(\theta_2,y)\Big]\,dy \\
&\qquad\qquad \text{(by using arguments similar to (2.10))} \\
&\le\ \frac{1}{4}\,(\theta_2-\theta_1)^2\,\sup_{\theta,y} J(\theta,y) \\
&=\ \frac{(u_2-u_1)^2}{4\,n_T}\,\sup_{\theta,y} J(\theta,y) \\
&\le\ \frac{C_0}{4}\,(u_2-u_1)^2 \qquad \text{(by (A3)(iii))}\ .
\end{aligned}
$$

This completes the proof of (a).

Let us now prove (b). By Hölder inequality,

$$
\begin{aligned}
E_\theta\Big[Z_T^{1/2}(u)\Big] =& \\
&= E_\theta\left[\exp\left\{\frac{1}{2}\int_0^T g_t(u)\,dW_t - \frac{1}{4}\int_0^T g_t^2(u)\,dt\right\}\right] \\
&= E_\theta\left[\exp\left\{\frac{1}{2}\int_0^T g_t(u)\,dW_t - \frac{1}{6}\int_0^T \big(g_t(u)\big)^2 dt\right\}\exp\left\{-\frac{1}{12}\int_0^T \big(g_t(u)\big)^2 dt\right\}\right] \\
&\le \left\{E_\theta\left[\exp\left\{\frac{1}{2}\int_0^T g_t(u)\,dW_t - \frac{1}{6}\int_0^T \big(g_t(u)\big)^2 dt\right\}\right]^{4/3}\right\}^{3/4} \qquad (2.14)\\
&\quad\times\left\{E_\theta\left[\exp\left\{-\frac{1}{12}\int_0^T \big(g_t(u)\big)^2 dt\right\}\right]^4\right\}^{1/4} \\
&\le \left[E_\theta\,\exp\left\{\frac{2}{3}\int_0^T g_t(u)\,dW_t - \frac{2}{9}\int_0^T \big(g_t^2(u)\big)\,dt\right\}\right]^{3/4} \\
&\quad\times\left[E_\theta\,\exp\left\{-\frac{1}{3}\int_0^T \big(g_t(u)\big)^2 dt\right\}\right]^{1/4}.
\end{aligned}
$$

Assumption (A5) implies that

$$
\begin{aligned}
E\,\exp\left\{-\frac{1}{3}\int_0^T \big(g_t(u)\big)^2 dt\right\} =& \\
&= E\,\exp\left\{-\frac{1}{3}\int_0^T\Big[f\big(\theta+u n_T^{-1/2},t,X_t\big)-f(\theta,t,X_t)^2\Big]^2 dt\right\} \qquad (2.15)\\
&\le \exp\big(-C\,|u|^\gamma\big)\ .
\end{aligned}
$$

On the other hand, from Gikhman and Skorohod (1972, p. 82)

$$E_\theta\left[\exp\left\{\int_0^T \frac{2}{3}\, g_t(u)\, dW_t - \frac{1}{2}\int_0^T \Big(\frac{2}{3}\, g_t(u)\Big)^2 dt\right\}\right] \le 1 \;. \tag{2.16}$$

Combination of (2.14)–(2.16) completes the proof of (b). □

Lemma 2.4 (Michel and Pfanzagl (1971))**.** *Let Y and Z be two random variables on some probability space with $P(Z>0)=1$. Then for all $\epsilon>0$, we have*

$$\sup_{x\in\mathbb{R}}\left|P\left\{\frac{Y}{Z}\le x\right\} - \Phi(x)\right| \;=\; \sup_{x\in\mathbb{R}}\Big|P\big\{Y\le x\big\} - \Phi(x)\Big| \;+\; P\big\{|Z-1|>\epsilon\big\} \;+\; \epsilon \;.$$

The following is the generalization of the above lemma from non-random η to random η.

Lemma 2.5 (Oblakova (1989))**.** *Let Y, Z and η be three random variables on some probability space with $P(Z>0)=1$, and η is a positive random variable with $P\big\{0<\eta^2<\infty\big\}=1$, $E(\eta^{-1})<\infty$. Then for all $\epsilon>0$, we have*

$$\sup_{x\in\mathbb{R}}\left|P\left\{\frac{Y}{Z}\le x\right\} - \Phi(x)\right| = E\sup_{x\in\mathbb{R}}\Big|P\big\{Y\le x\,|\,\mathcal{G}\big\} - \widetilde{\Phi}(x)\Big| + 2P\big\{|Z-\eta|>\epsilon\big\} + \epsilon E(\eta^{-1})\;.$$

where $\widetilde{\Phi}(x) = P(\zeta\eta\le x\,|\,\eta)$, $\mathcal{G}=\sigma(\eta)\subset\mathcal{F}_0$ and ζ is $\mathcal{N}(0,1)$ random variable independent of η.

3. MAIN RESULTS

We obtain the following large deviations upper bound for the MLE.

Theorem 3.1. *Under the assumptions (A1)–(A4), for $\rho>0$, we have*

$$\sup_{\theta\in\Theta} P_\theta^T\Big\{n_T^{1/2}\,|\theta_T-\theta|\ge\rho\Big\} \;\le\; B\exp\big(-b|\rho|^\gamma\big)$$

for some positive constants b and B independent of ρ and T.

Proof: Let

$$S_T := \Big\{u\colon\; \theta + u\, n_T^{-1/2}\in\Theta\Big\}\;,$$

$$
\begin{aligned}
P_\theta^T\left\{n_T^{1/2}|\theta_T-\theta|>\rho\right\} &= P_\theta^T\left\{|\theta_T-\theta|>\rho\, n_T^{-1/2}\right\}\\
&\le P_\theta^T\left\{\sup_{\substack{|u|\ge\rho\\ u\in S_T}} L_T\big(\theta+uT^{-1/2}\big)\ge L_T(\theta)\right\}\\
&= P_\theta^T\left\{\sup_{|u|\ge\rho}\frac{L_T\big(\theta+uT^{-1/2}\big)}{L_T(\theta)}\ge 1\right\} \qquad (3.1)\\
&= P_\theta^T\left\{\sup_{|u|\ge\rho} Z_T(u)\ge 1\right\}\\
&\le \sum_{r=0}^{\infty} P_\theta^T\left\{\sup_{u\in\Gamma_r} Z_T(u)\ge 1\right\},
\end{aligned}
$$

where $\Gamma_r=[\rho+r,\rho+r+1]$. Applying Lemma 2.1 with $\zeta(u)=Z_T^{1/2}(u)$, we obtain from Lemma 2.3 that there exists a constant $B>0$ such that

$$
\sup_{\theta\in\Theta} E_\theta^T\left\{\sup_{\substack{|u_1-u_2|\le h\\ |u_1|,|u_2|\le l}}\left[Z_T^{1/2}(u_1)-Z_T^{1/2}(u_2)\right]\right\}\le B\,l^{1/2}\,h^{1/2}\,\log h^{-1}\,. \qquad (3.2)
$$

Divide Γ_r into subintervals of length at most $h>0$. The number n of subintervals is clearly less than or equal to $[\frac{1}{h}]+1$. Let $\Gamma_r^{(j)}$, $1\le j\in n$ be the subintervals chosen. Choose $u_j\in\Gamma_r^{(j)}$. Then

$$
\begin{aligned}
&P_\theta^T\left[\sup_{u\in\Gamma_r} Z_T(u)\ge 1\right]\le\\
&\quad\le \sum_{j=1}^{n} P_\theta^T\left[Z_T^{1/2}(u_j)\ge\frac12\right]+P_\theta^T\left\{\sup_{\substack{|u-v|\le h\\ |u|,|v|\le\rho+r+1}}\left|Z_T^{1/2}(u)-Z_T^{1/2}(v)\right|\ge\frac12\right\}\\
&\quad\le 2\sum_{j=1}^{n} E_\theta^T\left[Z_T^{1/2}(u_j)\right]+2\,B(\rho+r+1)^{1/2}\,h^{1/2}\,\log(h^{-1})\\
&\qquad\qquad\qquad\qquad\qquad\qquad\text{(by Markov inequality and (3.2))}\\
&\quad\le 2\,C\sum_{j=1}^{n}\exp\big(-C\,|u_j|^{\gamma}\big)+2\,B(\rho+r+1)^{1/2}\,h^{1/2}\,\log(h^{-1})\quad\text{(by Lemma 2.2)}\\
&\quad\le 2\,C\left(\left[\frac1h\right]+1\right)\exp\big\{-C(\rho+r)^{\gamma}\big\}+2\,B(\rho+r+1)^{1/2}\,h^{1/2}\,\log(h^{-1})\,.
\end{aligned}
$$

Let us now choose $h=\exp\left\{\frac{-C(\rho+r)^{\gamma}}{2}\right\}$. Then

$$
\begin{aligned}
\sup_{\theta\in\Theta} P_\theta^T\left\{\sup_{u>\rho} Z_T(u)\ge 1\right\} &\le B\sum_{r=0}^{\infty}(\rho+r+1)^{1/2}\exp\left\{\frac{-C(\rho+r)^{\gamma}}{4}\right\} \qquad (3.3)\\
&\le B\exp(-b\rho^{\gamma})\,,
\end{aligned}
$$

where B and b are positive generic constants independent of ρ and T. Similarly it can be shown that

$$
\sup_{\theta\in\Theta} P_\theta^T\left[\sup_{u<-\rho} Z_T(u)\ge 1\right]\le B\exp(-b\rho^{\gamma})\,. \qquad (3.4)
$$

Combining (3.3) and (3.4), we obtain

$$\sup_{\theta\in\Theta} P_\theta^T \left[\sup_{|u|>\rho} Z_T(u) \geq 1 \right] \leq B \exp(-b\rho^\gamma) \ . \tag{3.5}$$

The theorem follows from (3.2) and (3.5). □

By substituting $\rho = n_T^{1/2}\epsilon$ in Theorem 3.1, the following result is obtained.

Corollary 3.1. *Under the conditions of Theorem 3.1, for arbitrary $\epsilon > 0$ and all $T > 0$, we have*

$$\sup_{\theta\in\Theta} P_\theta^T \left\{ |\theta_T - \theta| > \epsilon \right\} \leq B \exp(-b\, n_T \epsilon^\gamma)$$

where B and b are positive constants independent of ϵ and T.

We obtain the following large deviations bound for the Bayes estimator $\widetilde{\theta}_T$.

Theorem 3.2. *Suppose (A1)–(A4) and (B1)–(B5) hold. For $\rho > 0$, the Bayes estimator $\widetilde{\theta}_T$ with respect to the prior $\lambda(\cdot)$ and a loss function $l(\cdot,\cdot)$ with $l(u) = l(|u|)$ satisfies*

$$\sup_{\theta\in\Theta} P_\theta^T \left\{ \sqrt{T} \left| \widetilde{\theta}_T - \theta \right| \geq \rho \right\} \leq B \exp(-b\rho^2)$$

for some positive constants B and b independent of ρ and T.

Proof: Using Lemma 2.3, conditions (1), (2) and (3) of Lemma 2.2 are satisfied with $\alpha = 2$ and $g(u) = u^2$. Hence the result follows from Lemma 2.2. □

Corollary 3.2. *Under the conditions of Theorem 3.3, for arbitrary $\epsilon > 0$ and all $T > 0$, we have*

$$\sup_{\theta\in\Theta} P_\theta^T \left\{ \left| \widetilde{\theta}_T - \theta \right| > \epsilon \right\} \leq B \exp(-C\, T \epsilon^2) \ .$$

As another application of Theorem 3.3 we obtain the following result.

Theorem 3.3. *Under the assumptions (A1)–(A4), for all N, we have for the Bayes estimator $\widetilde{\theta}_T$ with respect to the prior $\lambda(\cdot)$ and loss function $\psi(\cdot,\cdot)$ satisfying the conditions (B1)–(B5),*

$$\lim_{\substack{H\to\infty \\ T\to\infty}} H^N \sup_{\theta\epsilon\Theta} P_\theta^T \left\{ \sqrt{T} \left| \widetilde{\theta}_T - \theta \right| > H \right\} = 0 \ .$$

We establish the following Berry–Esseen type inequality for the MLE.

Theorem 3.4. *Under the assumptions (A2), (A3) and (A5),*

$$\sup_{x\in\mathbb{R}}\left|P_\theta\Big\{I_T^{-1/2}(\theta)\,(\theta_T-\theta)\le x\Big\}-\Phi(x)\right|\le$$
$$\le E_\theta^{1/3}\left|\frac{I_T(\theta)}{m_T}-\eta(\theta)\right|+P_\theta\left\{\left|\frac{H_T(\theta)}{I_T}-1\right|>\frac{\epsilon_T}{2}\right\}+P_\theta\left\{\sup_{\theta\in\Theta}\big|I_T^{-1}(Q_T(\theta))\big|>\frac{\epsilon_T}{2\delta}\right\}$$
$$+\,C\exp\big(-b\,n_T\,\delta^2\big)\ .$$

for any $\delta>0$ and some $\epsilon_T\downarrow 0$ as $T\to\infty$ and $b>0$ is a constant independent of T.

Proof: Recall that $l'_t(\theta)\equiv U_T(\theta)$, $l''_T(\theta)\equiv H_T(\theta)$ and $l'''_T(\theta)\equiv Q_T(\theta)$.

By a Taylor expansion of $U_T(\theta)$ around θ, we have

$$0=U_T(\theta_T)=U_T(\theta)+(\theta_T-\theta)\,H_T(\overline{\theta}_T)\qquad\text{where }\ |\overline{\theta}_T-\theta|<|\theta_T-\theta|\ .$$

Hence

$$\begin{aligned}I_T^{1/2}(\theta)\,(\theta_T-\theta)&=-I_T^{1/2}(\theta)\,\frac{U_T(\theta)}{H_T(\overline{\theta}_T)}\\&=-\left(\frac{I_T(\theta)}{m_T}\right)^{1/2}\frac{m_T^{-1/2}\,U_T(\theta)}{m_T^{-1}\,H_T(\overline{\theta}_T)}\ .\end{aligned}$$

Thus by Lemma 2.3, we have

$$\begin{aligned}&\sup_{x\in\mathbb{R}}\left|P_\theta\Big\{I_T^{1/2}(\theta)\,(\theta_T-\theta)\le x\Big\}-\Phi(x)\right|=\\&\qquad=\sup_{x\in\mathbb{R}}\left|P_\theta\left\{\frac{-\left(\frac{I_T(\theta)}{m_T}\right)^{-1/2}m_T^{-1/2}\,U_T(\theta)}{\left(\frac{I_T(\theta)}{m_T}\right)^{-1}m_T^{-1}\,H_T(\overline{\theta}_T)}\le x\right\}-\Phi(x)\right|\\&\qquad=\sup_{x\in\mathbb{R}}\left|P_\theta\left\{\frac{-m_T^{-1/2}\,U_T(\theta)}{\left(\frac{I_T(\theta)}{m_T}\right)^{1/2}}\le x\right\}-\Phi(x)\right|\\&\qquad\quad+P_\theta\Big\{\big|I_T^{-1}(\theta)\,H_T(\overline{\theta}_T)-1\big|>\epsilon_T\Big\}+\epsilon_T\\&\qquad=:\ J_1+J_2+\epsilon_T\ .\end{aligned}$$

Let $M_T(\theta)=-m_T^{-1/2}\,U_T(\theta)=m_T^{-1/2}\int_0^T f'(\theta,t,X_t)\,dW_t$, a normalized continuous martingale with respect to $\mathcal{F}_T$ and $\langle M(\theta)\rangle_T=m_T^{-1}I_T(\theta)=m_T^{-1}\int_0^T f'^2(\theta,t,X_t)\,dt$ be its corresponding increasing process. Let $\widetilde{\Phi}(x)=P\big(G\eta\le x|\eta\big)$, $G\sim\mathcal{N}(0,1)$

and $\mathcal{G} \equiv \sigma(\eta) \subset \mathcal{F}_0$. Then by Lemma 2.5

$$\begin{aligned}
J_1 &= \sup_{x\in\mathbb{R}} \left| P_\theta \left\{ \frac{-m_T^{-1/2}\, U_T(\theta)}{\left(\frac{I_T(\theta)}{m_T}\right)^{1/2}} \le x \right\} - \Phi(x) \right| \\
&= \sup_{x\in\mathbb{R}} \left| P_\theta \left\{ \frac{M_T(\theta)}{\sqrt{\langle M(\theta)\rangle_T}} \le x \right\} - \Phi(x) \right| \\
&\le E_\theta \sup_{x\in\mathbb{R}} \left| P_\theta\big(M_T(\theta) \le x \,|\, \mathcal{G}\big) - \widetilde{\Phi}(x) \right| \\
&\quad + 2\, P_\theta \left\{ \left| \sqrt{\langle M(\theta)\rangle_T} - \eta(\theta) \right| > \epsilon_T \right\} + \epsilon_T\, E\big(\eta^{-1}(\theta)\big) \\
&\le C_1\, E_\theta^{1/3} \left| \frac{I_T(\theta)}{m_T} - \eta(\theta) \right| \qquad \text{(by Lemma 2.5)}
\end{aligned}$$

where C_1 depends only on $E(\eta^{-1}(\theta))$. Further,

$$\begin{aligned}
J_2 &= P_\theta \left\{ \left| I_T^{-1} H_T(\overline{\theta}_T) - 1 \right| > \epsilon_T \right\} \\
&\le P_\theta \left\{ \left| I_T^{-1} \right| \left| H_T(\overline{\theta}_T) - H_T(\theta) \right| > \frac{\epsilon_T}{2} \right\} + P_\theta \left\{ \left| I_T^{-1} H_T(\theta) - 1 \right| > \frac{\epsilon_T}{2} \right\} \\
&= P_\theta \left\{ I_T^{-1} \left| (\overline{\theta}_T - \theta)\, Q_T(\theta_T^\star) \right| > \frac{\epsilon_T}{2} \right\} + P_\theta \left\{ \left| I_T^{-1} H_T(\theta) - 1 \right| > \frac{\epsilon_T}{2} \right\} \\
&\qquad \big(\text{where } \ |\theta_T^\star - \theta| < |\overline{\theta}_T - \theta|\big) \\
&\le P_\theta \left\{ \left| I_T^{-1}\big(Q_T(\theta_T^\star)\big) \right| > \frac{\epsilon_T}{2\,\delta} \right\} + P_\theta \left\{ \left| \overline{\theta}_T - \theta \right| > \delta \right\} + P_\theta \left\{ \left| I_T^{-1} H_T(\theta) - 1 \right| > \frac{\epsilon_T}{2} \right\} \\
&\le P_\theta \left\{ \left| I_T^{-1}\big(Q_T(\theta_T^\star)\big) \right| > \frac{\epsilon_T}{2\,\delta} \right\} + P_\theta \left\{ \left| \theta_T - \theta \right| > \delta \right\} + P_\theta \left\{ \left| I_T^{-1} H_T(\theta) - 1 \right| > \frac{\epsilon_T}{2} \right\} \\
&\le P_\theta \left\{ \sup_{\theta\in\Theta} \left| I_T^{-1}\big(Q_T(\theta)\big) \right| > \frac{\epsilon_T}{2\,\delta} \right\} + C \exp\big(-b\, n_T\, \delta^\gamma\big) + P_\theta \left\{ \left| I_T^{-1} H_T(\theta) - 1 \right| > \frac{\epsilon_T}{2} \right\} \\
&\qquad \text{(by Corollary 3.2)}\ .
\end{aligned}$$

This completes the proof of the theorem. □

Remark. We used the splitting technique developed by Michel and Pfanzagl (1971) for the i.i.d. case. The upper bound in the Berry–Esseen type inequality obtained here contains four terms. The first term is cube root of the absolute moment, the second and the third term are moderate deviations type probabilities of the second and the third derivatives of log-likelihood respectively, and the fourth term is decays exponentially. The bound is quite sharp as seen in the linear case in the following example.

4. NONHOMOGENEOUS ORNSTEIN–UHLENBECK PROCESS

We apply the Berry–Esseen results for the MLE in the nonhomogeneous Ornstein–Uhlenbeck process satisfying the stochastic differential equation

$$dX_t = \theta\, t\, X_t\, dt + dW_t\ , \qquad t \geq 0, \quad X_0 = 0 \tag{4.1}$$

where $\theta > 0$. Note that the solution is a nonstationary and nonergodic process. Here the MLE based on $\{X_t,\, 0 \leq t \leq T\}$ is given by

$$\theta_T = \frac{\int_0^T t\, X_t\, dX_t}{\int_0^T t^2 X_t^2\, dt}\ ,$$

and $I_T(\theta) = \int_0^T t^2 X_t^2\, dt$. Let us choose $m_T = \int_0^T t^2\, e^{\theta t^2} dt$. Note that $m_T^{1/2}(\theta)(\theta_T - \theta)$ converges to Cauchy distribution with parameters $(0,1)$ as $T \to \infty$. Here $\frac{I_T}{m_T} \to \Delta^2$ a.s. where Δ has $\mathcal{N}\big(0, (\frac{\pi}{4\theta})^{1/2}\big)$ distribution and $\eta^2(\theta) = \Delta^2$. Directly from the calculation of J_1 in Theorem 3.6, we have

$$\begin{aligned}\sup_{x\in\mathbb{R}} \left| P_\theta \left\{ I_T^{1/2}(\theta)\, (\theta_T - \theta) \leq x \right\} - \Phi(x) \right| &\leq E_\theta^{1/3} \left| \frac{\int_0^T t^2 X_t^2\, dt}{\int_0^T t^2\, e^{\theta t^2}\, dt} - \Delta^2 \right| \\ &\leq C\, T^{1/2} \exp\left(\frac{-\theta\, T^4}{12} \right) \\ &\leq C \exp\left(\frac{-\theta\, T^4}{24} \right) .\end{aligned}$$

This shows that rate of weak convergence can be faster in the nonergodic processes than in ergodic processes in which case the sharpest possible rate is $O(T^{-1/2})$.

Remarks.

(1) Levanony *et al.* (1994) (see also Trofimov (1982)) showed that, for large enough t, MLE θ_t is a continuous semimartingale satisfying a stochastic differential equation. One could use the Berry–Esseen bound for semimartingales (see, e.g., Liptser and Shiryayev (1982, 1989)) to obtain a Berry–Esseen bound for the MLE θ_t. However, it would not give sharp bounds. Hence we follow the method of Michel and Pfanzagl (1971)) developed for the independent observations case.

(2) Large deviations for M-estimator remains to be investigated.

(3) It would be interesting if one can improve the Berry–Esseen bound in the above example by applying the characteristic function technique used in Bishwal (2000a).

(4) Berry–Esseen type bounds for Bayes estimators remain open.

(5) Large deviations and Berry–Esseen results for diffusions based on discrete observations remains to be investigated which would be more interesting in view of applications in finance.

REFERENCES

[1] Bahadur, R.R.; Zabell, S.L. and Gupta, J.C. (1980). *Large deviations, tests and estimates.* In "Asymptotic Theory of Statistical Tests" (I.M. Chakravarti, Ed.), Academic Press, New York, 33–64.

[2] Basawa, I.V. and Scott, D.J. (1983). *Asymptotic Optimal Inference for Non-ergodic Models*, Lecture Notes in Statistics **17**, Springer, Berlin.

[3] Bishwal, J.P.N. (1999). Large deviations inequalities for the maximum likelihood estimator and the Bayes estimators in nonlinear stochastic differential equations, *Statist. Probab. Letters*, **43**(2), 207–215.

[4] Bishwal, J.P.N. (2000a). Sharp Berry–Esseen bound for the maximum likelihood estimator in the Ornstein–Uhlenbeck process, *Sankhyā Ser. A*, **62**(1), 1–10.

[5] Bishwal, J.P.N. (2000b). Rates of convergence of the posterior distributions and the Bayes estimators in the Ornstein–Uhlenbeck process, *Random Operators and Stochastic Equations*, **8**(1), 51–70.

[6] Bishwal, J.P.N. (2005). Uniform rate of weak convergence for the minimum contrast estimator in the Ornstein–Uhlenbeck process, *Methodology and Computing in Applied Probability* (to appear).

[7] Bishwal, J.P.N. (2007). *Parameter Estimation in Stochastic Differential Equations*, Lecture Notes in Mathematics, Vol. **1923**, Springer-Verlag.

[8] Bishwal, J.P.N. and Bose, A. (1995). Speed of convergence of the maximum likelihood estimator in the Ornstein–Uhlenbeck process, *Calcutta Statist. Assoc. Bulletin*, **45**, 245–251.

[9] Borkar, V. and Bagchi, A. (1982). Parameter estimation in continuous time stochastic processes, *Stochastics*, **8**, 193–212.

[10] Bose, A. (1986). Berry–Esseen bound for the maximum likelihood estimator in the Ornstein–Uhlenbeck process, *Sankhyā Ser. A*, **48**, 181–187.

[11] Dupuis, P. and Kushner, H.J. (1989). Stochastic approximations and large deviations: upper bounds and w.p.1 convergence, *SIAM J. Control Optim.*, **27**, 1108–11035.

[12] Florens, D. and Pham, H. (1999). Large deviations in estimation an Ornstein–Uhlenbeck model, *J. Appl. Probab.*, **36**, 60–77.

[13] Gikhman, I.I. and Skorohod, A.V. (1972). *Stochastic Differential Equations*, Springer, Berlin.

[14] Harison, V. (1992). *The Bernstein-von Mises theorem for a certain class of Gaussian Diffusion processes*, Publications du Service de Mathematiques, Vol. 6, Universite d'Antananarivo, 1–7.

[15] Ibragimov, I.A. and Khasminskii, R.Z. (1981). *Statistical Estimation: Asymptotic Theory*, Springer-Verlag, Berlin.

[16] Jeganathan, P. (1982). On the asymptotic theory of statistical inference when the limit of the log-likelihood ratios is mixed normal, *Sankhyā Ser. A*, **44**, 173–212.

[17] Jeganathan, P. (1995). Some aspects of asymptotic theory with applications to time series models, *Econometric Theory*, **11**, 818–887.

[18] Kallianpur, G. and Selukar, R.S. (1993). *Estimation of Hilbert space valued parameter by the method of sieves*. In "Statistics and Probability: A Raghu Raj Bahadur Festschrift" (J.K. Ghosh, S.K. Mitra, K.R. Parthasarathy and B.L.S. Prakasa Rao, Eds.), Wiley Eastern, 325–347.

[19] Kutoyants, Yu.A. (1978). Estimation of parameter of a diffusion type process, *Theor. Probab. Applic.*, **23**, 641–649.

[20] Kutoyants, Yu.A. (1984). *Parameter Estimation for Stochastic Processes* (translated and edited by B.L.S. Prakasa Rao), Heldermann-Verlag, Berlin.

[21] Kutoyants, Yu.A. (1994). *Identification of Dynamical Systems with Small Noise*, Kluwer, Dordrecht.

[22] Le Cam, L. and Yang, G.L. (1990). *Asymptotics in Statistics*, Springer-Verlag, New York.

[23] Levanony, D. (1994). Conditional tail probabilities in continuous-time martingale LLN with application to parameter estimation in diffusions, *Stochastic Process. Appl.*, **51**, 117–134.

[24] Levanony, D.; Shwartz, A. and Zeitouni, O. (1994). Recursive identification in continuous-time stochastic processes, *Stochastic Process Appl.*, **49**, 245–275.

[25] Liptser, R.S. and Shiryayev, A.N. (1977). *Statistics of Random Processes I*, Springer-Verlag, Berlin.

[26] Liptser, R.S. and Shiryayev, A.N. (1982). On the rate of convergence in the central limit theorem for semimartingales, *Theory Probab. Appl.*, **27**, 1–13.

[27] Liptser, R.S. and Shiryayev, A.N. (1989). *Theory of Martingales*, Kluwer, Doredrecht.

[28] Mishra, M.N. (1989). The Bernstein- von Mises theorem for a class of nonhomogeneous diffusion processes, *Statist. Decisions*, **7**, 153–165.

[29] Mishra, M.N. and Prakasa Rao, B.L.S. (1985a). Asymptotic study of maximum likelihood estmation for nonhomogeneous diffusion processes, *Statist. Decisions*, **3**, 193–203.

[30] Michel, R. and Pfanzagl, J. (1971). The accuracy of the normal approximation for minimum contrast estimate, *Zeit Wahr. Verw. Gebiete*, **18**, 73–84.

[31] Mishra, M.N. and Prakasa Rao, B.L.S. (1985b). On the Berry–Esseen theorem for maximum likelihood estimator for linear homogeneous diffusion processes, *Sankhyā, Ser. A*, **47**, 392–398.

[32] Oblakova, T.V. (1989). On the rate of convergence in the central limit theorem for stochastic integrals, *Russian Math. Surveys*, **44**(2), 289–290.

[33] Sakamoto, Y. and Yoshida, N. (1998). Asymptotic expansion of M-estimator over Wiener space, *Statistical Inference for Stochastic Processes*, **1**, 85–103.

[34] Schmetterer, L. (1974). *Introduction to Mathematical Statistics*, Springer-Verlag, Berlin.

[35] Trofimov, E.I. (1982). A stochastic equation for the maximum likelihood estimate of the parameter of a diffusion type process, *Russian Math. Surveys*, **37**(4), 221–222.

[36] Yoshida, N. (1990). Asymptotic behaviour of M-estimator and related random field for diffusion process, *Ann. Inst. Statist. Math.*, **42**, 221–251.

[37] Yoshida, N. (1997). Malliavin calculus and asymptotic expansions of martingales, *Probab. Theory Relat. Fields*, **109**, 301–342.

REVSTAT – Statistical Journal
Volume 5, Number 3, November 2007, 269–283

MEASURE OF DEPARTURE FROM EXTENDED MARGINAL HOMOGENEITY FOR SQUARE CONTINGENCY TABLES WITH ORDERED CATEGORIES

Authors: Kouji Yamamoto
– Graduate School of Sciences and Technology,
Tokyo University of Science,
Noda City, Chiba, 278-8510, Japan
yamamoto_kouji@ymail.plala.or.jp

Yukari Furuya
– Graduate School of Sciences and Technology,
Tokyo University of Science,
Noda City, Chiba, 278-8510, Japan

Sadao Tomizawa
– Department of Information Sciences,
Faculty of Science and Technology, Tokyo University of Science,
Noda City, Chiba, 278-8510, Japan
tomizawa@is.noda.tus.ac.jp

Received: April 2007 Revised: September 2007 Accepted: October 2007

Abstract:

- For the analysis of square contingency tables, Tomizawa and Makii (2001), and Tomizawa, Miyamoto and Ashihara (2003) considered the measures to represent the degree of departure from marginal homogeneity (MH). Tomizawa (1984) considered an extended marginal homogeneity (EMH) model for square tables with ordered categories. This paper proposes a measure to represent the degree of departure from EMH. The measure proposed is expressed by using the Cressie and Read's (1984) power-divergence or Patil and Taillie's (1982) diversity index. The measure would be useful for comparing the degree of departure from EMH in several tables. Examples are given.

Key-Words:

- *cumulative probability; Kullback–Leibler information; marginal homogeneity; measure; ordered category; power-divergence; Shannon entropy; square contingency table.*

AMS Subject Classification:

- 62H17.

1. INTRODUCTION

Consider an $R\times R$ square contingency table with the same row and column classifications. Let p_{ij} denote the probability that an observation will fall in the i-th row and j-th column of the table ($i=1,...,R$; $j=1,...,R$), and let X and Y denote the row and column variables, respectively. The marginal homogeneity (MH) model is defined by

$$\Pr(X=i) = \Pr(Y=i) \qquad \text{for} \;\; i=1,...,R \;,$$

namely

$$p_{i\cdot} = p_{\cdot i} \qquad \text{for} \;\; i=1,...,R \;,$$

where $p_{i\cdot} = \sum_{k=1}^{R} p_{ik}$ and $p_{\cdot i} = \sum_{k=1}^{R} p_{ki}$ (see, e.g., Stuart, 1955; Bhapkar, 1966; Bishop, Fienberg and Holland, 1975, p. 294).

Let

$$G_{1(i)} = \sum_{s=1}^{i} \sum_{t=i+1}^{R} p_{st} \qquad \bigl[= \Pr\bigl(X\leq i,\, Y\geq i+1\bigr)\bigr] \;,$$

and

$$G_{2(i)} = \sum_{s=i+1}^{R} \sum_{t=1}^{i} p_{st} \qquad \bigl[= \Pr\bigl(X\geq i+1,\, Y\leq i\bigr)\bigr] \;,$$

for $i=1,...,R-1$. By considering the difference between the cumulative marginal probabilities, $F_i^X - F_i^Y$ for $i=1,...,R-1$, where $F_i^X = \Pr(X\leq i)$ and $F_i^Y = \Pr(Y\leq i)$, we see that the MH model may also be expressed as

$$G_{1(i)} = G_{2(i)} \qquad \text{for} \;\; i=1,...,R-1 \;.$$

Namely, this states that the cumulative probability that an observation will fall in row category i or below and column category $i+1$ or above is equal to the cumulative probability that the observation falls in column category i or below and row category $i+1$ or above for $i=1,...,R-1$.

Tomizawa (1984, 1995a) considered the extended marginal homogeneity (EMH) model defined by

$$p_{i\cdot}^{(\delta)} = p_{\cdot i}^{(\delta)} \qquad \text{for} \;\; i=1,...,R \;,$$

where the parameter δ is unspecified and

$$p_{i\cdot}^{(\delta)} = \delta \sum_{t=1}^{i-1} p_{it} + \sum_{t=i}^{R} p_{it} \;, \qquad p_{\cdot i}^{(\delta)} = \sum_{s=1}^{i} p_{si} + \delta \sum_{s=i+1}^{R} p_{si} \;.$$

Consider the artificial probabilities in Table 1. We see that the EMH model holds with $\delta=2$ in Table 1. The EMH model may also be expressed as

$$G_{1(i)} = \delta\, G_{2(i)} \qquad \text{for} \;\; i=1,...,R-1 \;.$$

Table 1: Artificial probabilities having the structure of EMH with $\delta = 2$.

0.04	0.02	0.04	0.26
0.01	0.03	0.08	0.16
0.02	0.04	0.02	0.04
0.13	0.08	0.02	0.01

A special case of this model obtained by putting $\delta = 1$ is the MH model. This model indicates that the cumulative probability that an observation will fall in row category i or below and column category $i+1$ or above is δ times higher than the cumulative probability that the observation falls in column category i or below and row category $i+1$ or above for $i = 1, ..., R-1$. The EMH model may further be expressed as

$$G^*_{1(i)} = G^*_{2(i)} \qquad \text{for} \quad i = 1, ..., R-1 \ , \tag{1.1}$$

where

$$G^*_{1(i)} = G_{1(i)}/G_1 \ , \qquad G^*_{2(i)} = G_{2(i)}/G_2 \ ,$$

$$G_1 = \sum_{i=1}^{R-1} G_{1(i)} \ , \qquad G_2 = \sum_{i=1}^{R-1} G_{2(i)} \ .$$

Namely the EMH model indicates that there is a structure of symmetry between $\{G^*_{1(i)}\}$ and $\{G^*_{2(i)}\}$ for $i = 1, ..., R-1$.

For square contingency tables with *nominal* categories, Tomizawa (1995b) considered two kinds of measures to represent the degree of departure from MH, which are expressed by using the Shannon entropy and Gini concentration. Tomizawa and Makii (2001) considered a generalization of Tomizawa measures, which is expressed by using the Cressie and Read's (1984) power-divergence (or Patil and Taillie's (1982) diversity index). For square contingency tables with *ordered* categories, Tomizawa, Miyamoto and Ashihara (2003) considered a measure to represent the degree of departure from MH.

When the MH model does not hold, these measures would be useful for measuring the degree of departure from MH. When the EMH model does not hold, we are now interested in measuring the degree of departure from EMH (instead of that from MH).

The purpose of this paper is to propose a power-divergence type measure which represents the degree of departure from EMH for square contingency tables with ordered categories. In Section 2 we propose such a measure which is expressed as a function of $\{G^*_{1(i)}\}$ and $\{G^*_{2(i)}\}$. It would be useful for *comparing* the degree of departure from EMH in several tables with ordered categories.

2. MEASURE OF DEPARTURE FROM EXTENDED MARGINAL HOMOGENEITY

Assume that $G_1 > 0$, $G_2 > 0$ and $G_{1(i)} + G_{2(i)} > 0$ for $i = 1, ..., R-1$. Let

$$C_i = \frac{G^*_{1(i)} + G^*_{2(i)}}{2} \qquad \text{for} \quad i = 1, ..., R-1 \ .$$

Note that $\sum_{i=1}^{R-1} C_i = 1$. To represent the degree of departure from EMH, consider a measure defined by

$$\Gamma^{(\lambda)}_{\mathrm{EM}} = \frac{\lambda(\lambda+1)}{2\,(2^\lambda-1)} \Big[I^{(\lambda)}\Big(\big\{G^*_{1(i)}\big\}; \big\{C_i\big\}\Big) + I^{(\lambda)}\Big(\big\{G^*_{2(i)}\big\}; \big\{C_i\big\}\Big)\Big] \qquad \text{for} \quad \lambda > -1 \ ,$$

where

$$I^{(\lambda)}\big(\{a_i\}; \{b_i\}\big) = \frac{1}{\lambda(\lambda+1)} \sum_{i=1}^{R-1} a_i \left[\left(\frac{a_i}{b_i}\right)^\lambda - 1 \right] ,$$

and the value at $\lambda = 0$ is taken to be the limit as $\lambda \to 0$. Thus,

$$\Gamma^{(0)}_{\mathrm{EM}} = \frac{1}{2\log 2} \Big[I^{(0)}\Big(\big\{G^*_{1(i)}\big\}; \big\{C_i\big\}\Big) + I^{(0)}\Big(\big\{G^*_{2(i)}\big\}; \big\{C_i\big\}\Big)\Big] \ , \tag{2.1}$$

where

$$I^{(0)}\big(\{a_i\}; \{b_i\}\big) = \sum_{i=1}^{R-1} a_i \log\left(\frac{a_i}{b_i}\right) .$$

The $I^{(\lambda)}\big(\{a_i\}; \{b_i\}\big)$ is the power-divergence between $\{a_i\}$ and $\{b_i\}$, and especially $I^{(0)}\big(\{a_i\}; \{b_i\}\big)$ is the Kullback–Leibler information (KL) between them. For more details of the power-divergence $I^{(\lambda)}(\cdot;\cdot)$, see Cressie and Read (1984), and Read and Cressie (1988, p. 15). Note that a real value λ is chosen by user.

Let

$$G^c_{1(i)} = \frac{G^*_{1(i)}}{G^*_{1(i)} + G^*_{2(i)}} \ , \quad G^c_{2(i)} = \frac{G^*_{2(i)}}{G^*_{1(i)} + G^*_{2(i)}} \qquad \text{for} \quad i = 1, ..., R-1 \ .$$

Note that $\big\{G^c_{1(i)} + G^c_{2(i)} = 1\big\}$. The EMH model can be expressed as

$$G^c_{1(i)} = G^c_{2(i)} \left(= \frac{1}{2}\right) \qquad \text{for} \quad i = 1, ..., R-1 \ .$$

Then the measure $\Gamma^{(\lambda)}_{\mathrm{EM}}$ may be expressed as

$$\Gamma^{(\lambda)}_{\mathrm{EM}} = \frac{\lambda(\lambda+1)}{2^\lambda - 1} \sum_{i=1}^{R-1} C_i \, I^{(\lambda)}_i \left(\Big\{ G^c_{1(i)}, G^c_{2(i)} \Big\} ; \left\{ \frac{1}{2}, \frac{1}{2} \right\} \right) \qquad \text{for} \quad \lambda > -1 \ , \tag{2.2}$$

where

$$I_i^{(\lambda)}(\cdot;\cdot) \;=\; \frac{1}{\lambda(\lambda+1)}\left[G^c_{1(i)}\left\{\left(\frac{G^c_{1(i)}}{1/2}\right)^{\lambda}-1\right\}+G^c_{2(i)}\left\{\left(\frac{G^c_{2(i)}}{1/2}\right)^{\lambda}-1\right\}\right] ,$$

and the value at $\lambda=0$ is taken to be the limit as $\lambda\to 0$. Thus,

$$\Gamma^{(0)}_{\mathrm{EM}} \;=\; \frac{1}{\log 2}\sum_{i=1}^{R-1} C_i\, I_i^{(0)}\left(\left\{G^c_{1(i)}, G^c_{2(i)}\right\};\left\{\frac{1}{2},\frac{1}{2}\right\}\right) ,$$

where

$$I_i^{(0)}(\cdot;\cdot) \;=\; G^c_{1(i)}\log\left(\frac{G^c_{1(i)}}{1/2}\right)+G^c_{2(i)}\log\left(\frac{G^c_{2(i)}}{1/2}\right) .$$

Therefore, $\Gamma^{(\lambda)}_{\mathrm{EM}}$ in equation (2.2) would represent, essentially, the weighted sum of the power-divergence $I_i^{(\lambda)}\left(\left\{G^c_{1(i)}, G^c_{2(i)}\right\};\{1/2,1/2\}\right)$.

Moreover, $\Gamma^{(\lambda)}_{\mathrm{EM}}$ may be expressed as

$$\text{(2.3)}\qquad \Gamma^{(\lambda)}_{\mathrm{EM}} \;=\; 1-\frac{\lambda\, 2^{\lambda}}{2^{\lambda}-1}\sum_{i=1}^{R-1} C_i\, H_i^{(\lambda)}\left(\left\{G^c_{1(i)}, G^c_{2(i)}\right\}\right) \qquad \text{for}\ \ \lambda>-1 ,$$

where

$$H_i^{(\lambda)}(\cdot) \;=\; \frac{1}{\lambda}\left[1-\left(G^c_{1(i)}\right)^{\lambda+1}-\left(G^c_{2(i)}\right)^{\lambda+1}\right] ,$$

and the value at $\lambda=0$ is taken to be the limit as $\lambda\to 0$. Thus,

$$\Gamma^{(0)}_{\mathrm{EM}} \;=\; 1-\frac{1}{\log 2}\sum_{i=1}^{R-1} C_i\, H_i^{(0)}\left(\left\{G^c_{1(i)}, G^c_{2(i)}\right\}\right) ,$$

where

$$H_i^{(0)}(\cdot) \;=\; -G^c_{1(i)}\log G^c_{1(i)} - G^c_{2(i)}\log G^c_{2(i)} .$$

Note that $H_i^{(\lambda)}\left(\left\{G^c_{1(i)}, G^c_{2(i)}\right\}\right)$ is Patil and Taillie's (1982) diversity index for $\left\{G^c_{1(i)}, G^c_{2(i)}\right\}$, which includes the Shannon entropy (when $\lambda=0$) and the Gini concentration (when $\lambda=1$) in special cases. Therefore, $\Gamma^{(\lambda)}_{\mathrm{EM}}$ in equation (2.3) would represent essentially the weighted sum of the diversity index $H_i^{(\lambda)}\left(\left\{G^c_{1(i)}, G^c_{2(i)}\right\}\right)$.

Noting that $I_i^{(\lambda)}\left(\left\{G^c_{1(i)}, G^c_{2(i)}\right\};\{1/2,1/2\}\right)\ge 0$ and $H_i^{(\lambda)}\left(\left\{G^c_{1(i)}, G^c_{2(i)}\right\}\right)\ge 0$, we see that the measure $\Gamma^{(\lambda)}_{\mathrm{EM}}$ must lie between 0 and 1. Also, for each λ (>-1),

- **(i)** there is a structure of EMH in the $R\times R$ table if and only if $\Gamma^{(\lambda)}_{\mathrm{EM}}=0$, and
- **(ii)** the degree of departure from EMH is the largest in the sense that $G^c_{1(i)}=0$ (then $G^c_{2(i)}=1$) or $G^c_{2(i)}=0$ (then $G^c_{1(i)}=1$) [namely, $G^*_{1(i)}=0$ (then $G^*_{2(i)}>0$) or $G^*_{2(i)}=0$ (then $G^*_{1(i)}>0$)] for $i=1,\dots,R-1$; if and only if $\Gamma^{(\lambda)}_{\mathrm{EM}}=1$.

Note that $\Gamma^{(\lambda)}_{\mathrm{EM}} = 1$ indicates that $G^*_{1(i)}/G^*_{2(i)} = \infty$ for some i and $G^*_{1(i)}/G^*_{2(i)} = 0$ for the other i, and therefore it seems appropriate to consider that then the degree of departure from EMH (i.e., from $G^*_{1(i)}/G^*_{2(i)} = 1$ for $i = 1, ..., R-1$) is largest. In addition, according to the weighted sum of the power-divergence or the weighted sum of the Patil and Taillie's diversity index, the degree increases as the value of $\Gamma^{(\lambda)}_{\mathrm{EM}}$ increases.

3. APPROXIMATE CONFIDENCE INTERVAL FOR MEASURE

Let n_{ij} denote the observed frequency in the i-th row and j-th column of the table ($i = 1, ..., R$; $j = 1, ..., R$). Assuming that a multinomial distribution applies to the $R \times R$ table, we shall consider an approximate standard error and large-sample confidence interval for $\Gamma^{(\lambda)}_{\mathrm{EM}}$ using the delta method, descriptions of which are given by Bishop *et al.* (1975, Sec. 14.6) and Agresti (1990, Sec. 12.1). The sample version of $\Gamma^{(\lambda)}_{\mathrm{EM}}$, i.e., $\hat{\Gamma}^{(\lambda)}_{\mathrm{EM}}$, is given by $\Gamma^{(\lambda)}_{\mathrm{EM}}$ with $\{p_{ij}\}$ replaced by $\{\hat{p}_{ij}\}$, where $\hat{p}_{ij} = n_{ij}/n$ and $n = \sum\sum n_{ij}$. Using the delta method, $\sqrt{n}\big(\hat{\Gamma}^{(\lambda)}_{\mathrm{EM}} - \Gamma^{(\lambda)}_{\mathrm{EM}}\big)$ has asymptotically (as $n \to \infty$) a normal distribution with mean zero and variance,

$$\sigma^2\Big[\Gamma^{(\lambda)}_{\mathrm{EM}}\Big] = \sum_{k=1}^{R-1}\sum_{l=k+1}^{R}\Big[p_{kl}\big(\omega^{(\lambda)}_{1(kl)}\big)^2 + p_{lk}\big(\omega^{(\lambda)}_{2(kl)}\big)^2\Big] ,$$

where for $\lambda > -1$, $\lambda \neq 0$; $t = 1, 2$,

$$\omega^{(\lambda)}_{t(kl)} = \frac{2^\lambda}{2\,(2^\lambda - 1)\,G_t}\left[\sum_{i=k}^{l-1}\Delta^{(\lambda)}_{t(i)} - (l-k)\sum_{i=1}^{R-1} G^*_{t(i)}\,\Delta^{(\lambda)}_{t(i)}\right] ,$$

$$\Delta^{(\lambda)}_{1(i)} = \big(G^c_{1(i)}\big)^\lambda + \lambda\Big\{\big(G^c_{1(i)}\big)^\lambda - \big(G^c_{2(i)}\big)^\lambda\Big\}\,G^c_{2(i)} ,$$
$$\Delta^{(\lambda)}_{2(i)} = \big(G^c_{2(i)}\big)^\lambda + \lambda\Big\{\big(G^c_{2(i)}\big)^\lambda - \big(G^c_{1(i)}\big)^\lambda\Big\}\,G^c_{1(i)} ;$$

and for $\lambda = 0$; $t = 1, 2$,

$$\omega^{(0)}_{t(kl)} = \frac{1}{2\,(\log 2)\,G_t}\left[\sum_{i=k}^{l-1}\log\big(G^c_{t(i)}\big) - (l-k)\sum_{i=1}^{R-1} G^*_{t(i)}\log\big(G^c_{t(i)}\big)\right] .$$

We note that the asymptotic distribution of $\sqrt{n}\,\big(\hat{\Gamma}^{(\lambda)}_{\mathrm{EM}} - \Gamma^{(\lambda)}_{\mathrm{EM}}\big)$ is not applicable when $\Gamma^{(\lambda)}_{\mathrm{EM}} = 0$ and $\Gamma^{(\lambda)}_{\mathrm{EM}} = 1$ because then $\sigma^2\big[\Gamma^{(\lambda)}_{\mathrm{EM}}\big] = 0$. Let $\hat{\sigma}^2\big[\Gamma^{(\lambda)}_{\mathrm{EM}}\big]$ denote $\sigma^2\big[\Gamma^{(\lambda)}_{\mathrm{EM}}\big]$ with $\{p_{ij}\}$ replaced by $\{\hat{p}_{ij}\}$. Then $\hat{\sigma}\big[\Gamma^{(\lambda)}_{\mathrm{EM}}\big]/\sqrt{n}$ is an estimated approximate standard error for $\hat{\Gamma}^{(\lambda)}_{\mathrm{EM}}$, and $\hat{\Gamma}^{(\lambda)}_{\mathrm{EM}} \pm z_{p/2}\,\hat{\sigma}\big[\Gamma^{(\lambda)}_{\mathrm{EM}}\big]/\sqrt{n}$ is an approximate $100\,(1-p)$ percent confidence interval for $\Gamma^{(\lambda)}_{\mathrm{EM}}$, where $z_{p/2}$ is the percentage point from the standard normal distribution corresponding to a two-tail probability equal to p.

4. EXAMPLES

Consider the data in Table 2, taken from Tominaga (1979, p. 53). These data describe the cross-classification of father's and son's occupational status categories in Japan which were examined in 1955, 1965 and 1975.

Table 2: Occupational status for Japanese father-son pairs; from Tominaga (1979, p. 53).

Father's status	Son's status								Total
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	
(a) Examined in 1955									
(1)	36	4	14	7	8	2	3	8	82
(2)	20	20	27	24	11	11	2	11	126
(3)	9	6	23	12	9	5	3	16	83
(4)	15	14	39	81	17	16	11	15	208
(5)	6	7	22	13	72	20	6	13	159
(6)	3	2	5	12	18	19	9	7	75
(7)	5	3	10	11	21	15	38	25	128
(8)	39	30	76	80	69	52	45	614	1005
Total	133	86	216	240	225	140	117	709	1866
(b) Examined in 1965									
(1)	27	10	16	3	6	6	1	2	71
(2)	15	38	30	20	8	4	3	7	125
(3)	13	17	32	17	7	16	6	5	113
(4)	12	36	40	132	22	30	13	6	291
(5)	8	22	38	41	91	42	22	9	273
(6)	2	2	7	12	13	16	3	2	57
(7)	3	2	11	11	13	26	30	6	102
(8)	38	44	95	101	132	114	60	309	893
Total	118	171	269	337	292	254	138	346	1925
(c) Examined in 1975									
(1)	44	18	28	8	6	8	1	5	118
(2)	15	50	45	20	18	17	4	7	176
(3)	18	25	47	30	24	18	5	7	174
(4)	16	27	53	77	40	29	9	6	257
(5)	18	25	42	31	122	43	17	13	311
(6)	12	15	21	15	36	33	3	8	143
(7)	3	5	8	7	26	21	9	3	82
(8)	44	65	114	92	184	195	58	325	1077
Total	170	230	358	280	456	364	106	374	2338

Note: Status (1) is Professional, (2) Managers, (3) Clerical, (4) Sales, (5) Skilled manual, (6) Semiskilled manual, (7) Unskilled manual and (8) Farmers.

Since the confidence intervals for $\Gamma_{\text{EM}}^{(\lambda)}$ applied to the data in Tables 2a, 2b and 2c do not include zero for all λ (see Table 3), these would indicate that there is not a structure of EMH in each table.

Table 3: Estimate of $\Gamma_{\text{EM}}^{(\lambda)}$, estimated approximate standard error for $\hat{\Gamma}_{\text{EM}}^{(\lambda)}$, and approximate 95% confidence interval for $\Gamma_{\text{EM}}^{(\lambda)}$, applied to Tables 2a, 2b and 2c.

Values of λ	Estimated measure	Standard error	Confidence interval
	(a) For Table 2a		
−0.5	0.017	0.004	(0.009, 0.024)
0	0.028	0.006	(0.016, 0.040)
0.5	0.035	0.008	(0.019, 0.050)
1.0	0.038	0.009	(0.021, 0.055)
1.5	0.039	0.009	(0.022, 0.056)
2.0	0.038	0.009	(0.021, 0.055)
2.5	0.036	0.008	(0.020, 0.052)
	(b) For Table 2b		
−0.5	0.043	0.006	(0.031, 0.055)
0	0.070	0.009	(0.051, 0.088)
0.5	0.085	0.011	(0.063, 0.107)
1.0	0.093	0.012	(0.069, 0.116)
1.5	0.095	0.012	(0.071, 0.118)
2.0	0.093	0.012	(0.069, 0.116)
2.5	0.088	0.012	(0.066, 0.111)
	(c) For Table 2c		
−0.5	0.053	0.007	(0.040, 0.066)
0	0.086	0.010	(0.066, 0.106)
0.5	0.105	0.012	(0.081, 0.129)
1.0	0.114	0.013	(0.089, 0.139)
1.5	0.116	0.013	(0.091, 0.142)
2.0	0.114	0.013	(0.089, 0.139)
2.5	0.109	0.012	(0.084, 0.133)

When the degrees of departure from EMH in Tables 2a, 2b and 2c are compared using the confidence interval for $\Gamma_{\text{EM}}^{(\lambda)}$, it is greater in Tables 2b and 2c than in Table 2a. However, the comparison between Tables 2b and 2c may be impossible, because the values in the confidence interval for Table 2b are not always greater than the values in the confidence interval for Table 2c.

We shall investigate the degree of departure from EMH in more details. For instance, when $\lambda = 1$, the estimated measure $\hat{\Gamma}_{\text{EM}}^{(1)}$ equals 0.038 for Table 2a, 0.093 for Table 2b, and 0.114 for Table 2c (see Table 3). Thus,

(i) for Table 2a, the degree of departure from EMH is estimated to be 3.8 percent of the maximum degree of departure from EMH,

(ii) for Table 2b, it is estimated to be 9.3 percent of the maximum degree of departure from EMH,

and

(iii) for Table 2c, it is estimated to be 11.4 percent of the maximum degree of departure from EMH.

Note: Let $W^{(\lambda)}$ $(-\infty < \lambda < \infty)$ denote the power-divergence statistic for testing goodness-of-fit of the EMH model with $R-2$ degrees of freedom. [See Appendix for $W^{(\lambda)}$, and see Cressie and Read (1984) and Read and Cressie (1988, p.15) for details of the power-divergence test statistic.] In particular, $W^{(0)}$ and $W^{(1)}$ are the likelihood ratio and Pearson's chi-squared statistics, respectively. Table 4 gives the values of $W^{(\lambda)}$ applied to the data in Tables 2a, 2b, and 2c. These data fit the EMH model very poorly.

Table 4: The values of power-divergence statistic $W^{(\lambda)}$ (with 6 degrees of freedom) for testing goodness-of-fit of the EMH model, applied to Tables 2a, 2b and 2c.

Values of λ	For Table 2a	For Table 2b	For Table 2c
−0.5	118.52	300.36	333.41
0	116.76	231.58	280.73
0.5	117.38	200.39	252.56
1.0	120.39	186.77	239.04
1.5	125.95	183.14	235.42
2.0	134.42	186.48	239.54
2.5	146.33	195.69	250.66

5. CONCLUDING REMARKS

The measure $\Gamma_{\mathrm{EM}}^{(\lambda)}$ always ranges between 0 and 1 independent of the dimension R and sample size n. Therefore, $\Gamma_{\mathrm{EM}}^{(\lambda)}$ may be useful for *comparing* the degree of departure from EMH in several tables.

Consider the artificial data in Table 5. Table 6 gives the values of $W^{(\lambda)}$ (with 2 degrees of freedom) for testing goodness-of-fit of the EMH model applied to these data. Compare the values of $W^{(\lambda)}$ for Tables 5a and 5b. From $W^{(\lambda)}$ with any fixed λ, we see that the EMH model fits the data in Table 5a worse than the data in Table 5b (see Table 6). In contrast, for any fixed λ (> -1), the value of $\hat{\Gamma}_{\mathrm{EM}}^{(\lambda)}$ is less for Table 5a than for Table 5b (see Table 7). In terms of $\hat{G}_{1(i)}/\hat{G}_{2(i)}$, $i=1,2,3$ (see Table 5), it seems natural to conclude that the degree of departure from EMH is less for Table 5a than for Table 5b. Therefore $\hat{\Gamma}_{\mathrm{EM}}^{(\lambda)}$

Table 5: Artificial data.

(a) $n = 2829$

	(1)	(2)	(3)	(4)	Total
(1)	187	330	70	20	607
(2)	30	178	60	40	308
(3)	50	100	898	60	1108
(4)	70	20	10	706	806
Total	337	628	1038	826	2829

Note: $\frac{\hat{G}_{1(1)}}{\hat{G}_{2(1)}} = 2.80, \quad \frac{\hat{G}_{1(2)}}{\hat{G}_{2(2)}} = 0.79, \quad \frac{\hat{G}_{1(3)}}{\hat{G}_{2(3)}} = 1.20\,.$

(b) $n = 2654$

	(1)	(2)	(3)	(4)	Total
(1)	687	80	10	5	782
(2)	5	178	5	12	200
(3)	5	25	898	13	941
(4)	10	8	7	706	731
Total	707	291	920	736	2654

Note: $\frac{\hat{G}_{1(1)}}{\hat{G}_{2(1)}} = 4.75, \quad \frac{\hat{G}_{1(2)}}{\hat{G}_{2(2)}} = 0.67, \quad \frac{\hat{G}_{1(3)}}{\hat{G}_{2(3)}} = 1.20\,.$

(b) $n = 429$

	(1)	(2)	(3)	(4)	Total
(1)	68	80	10	5	163
(2)	5	17	5	12	39
(3)	5	25	89	13	132
(4)	10	8	7	70	95
Total	88	130	111	100	429

Note: $\frac{\hat{G}_{1(1)}}{\hat{G}_{2(1)}} = 4.75, \quad \frac{\hat{G}_{1(2)}}{\hat{G}_{2(2)}} = 0.67, \quad \frac{\hat{G}_{1(3)}}{\hat{G}_{2(3)}} = 1.20\,.$

may be preferable to $W^{(\lambda)}$ for *comparing* the degree of departure from EMH in several tables. It may seem, to many readers, that $W^{(\lambda)}/n$ (for a given λ) is also a reasonable measure for representing the degree of departure from EMH. However, it does not seem to us that $W^{(\lambda)}/n$ is a reasonable measure. For example, consider the artificial data in Tables 5b and 5c. The values of $W^{(\lambda)}/n$ are, for example, when $\lambda = 0$ ($\lambda = 1$), $W^{(0)}/n = 0.024$ ($W^{(1)}/n = 0.022$) for Table 5b, and $W^{(0)}/n = 0.147$ ($W^{(1)}/n = 0.138$) for Table 5c. Therefore the value of $W^{(\lambda)}/n$ is less for Table 5b than for Table 5c. On the other side, for any fixed λ (> -1), the value of $\hat{\Gamma}^{(\lambda)}_{\mathrm{EM}}$ for Table 5b is theoretically identical to that for Table 5c (see Table 7). In addition, $\hat{G}_{1(i)}/\hat{G}_{2(i)}$, $i=1,2,3$, for Table 5b is identical to that for Table 5c (see Table 5). So it seems natural to conclude that the degree of departure from EMH for Table 5b is equal to that for Table 5c. Therefore $\hat{\Gamma}^{(\lambda)}_{\mathrm{EM}}$ may also be preferable to $W^{(\lambda)}/n$ for *comparing* the degree of departure from EMH in several tables.

Table 6: The values of $W^{(\lambda)}$ (with 2 degrees of freedom) for testing goodness-of-fit of the EMH model, applied to Tables 5a, 5b and 5c.

Values of λ	For Table 5a	For Table 5b	For Table 5c
−0.5	194.43	69.35	69.35
0	182.76	63.25	63.25
0.5	175.25	60.04	60.04
1.0	171.21	59.04	59.04
1.5	170.17	59.93	59.93
2.0	171.89	62.61	62.61
2.5	176.28	67.17	67.17

Table 7: The values of $\hat{\Gamma}^{(\lambda)}_{\mathrm{EM}}$ applied to Tables 5a, 5b and 5c.

Values of λ	For Table 5a	For Table 5b	For Table 5c
−0.5	0.034	0.076	0.076
0	0.057	0.125	0.125
0.5	0.071	0.153	0.153
1.0	0.078	0.167	0.167
1.5	0.080	0.171	0.171
2.0	0.078	0.167	0.167
2.5	0.074	0.159	0.159

Since the EMH model is expressed as equation (1.1), we are interested in measuring how far $\{G^*_{1(i)}\}$ and $\{G^*_{2(i)}\}$ are distant from those with an EMH structure when the EMH model does not hold. The measure $\Gamma^{(\lambda)}_{\mathrm{EM}}$ is a function of $\{G^*_{1(i)}\}$ and $\{G^*_{2(i)}\}$. Since equation (1.1), it seems natural that the measure is expressed as a function of $\{G^*_{1(i)}\}$ and $\{G^*_{2(i)}\}$.

For the measure $\Gamma^{(\lambda)}_{\mathrm{EM}}$, the analyst may be interested in which value of λ is preferred for a given table. However, it seems difficult to discuss this. It seems to be important and safe that for *comparing* the degrees of departure from EMH in several tables, the analyst calculates the values of $\hat{\Gamma}^{(\lambda)}_{\mathrm{EM}}$ for various values of λ and discusses the degree of departure from EMH in terms of them (rather than calculating $\hat{\Gamma}^{(\lambda)}_{\mathrm{EM}}$ for *only* one specified value of λ).

Table 8: Artificial data.

(a) $n = 1895$ (sample size)

374	602	170	64
18	255	139	71
4	23	42	55
2	6	17	53

(b) $n = 5397$

81	444	632	726
646	178	498	6
787	288	68	762
72	105	17	87

Consider the artificial data in Tables 8a and 8b. Then we see from Table 9 that the value of $\hat{\Gamma}^{(0)}_{\rm EM}$ is less for Table 8a than for Table 8b, but the value of $\hat{\Gamma}^{(1)}_{\rm EM}$ is greater for Table 8a than for Table 8b. However, the differences are very slight in these cases. So, for these cases, we may conclude (by using $\hat{\Gamma}^{(\lambda)}_{\rm EM}$) that the departure from EMH for Table 8a is similar to that for Table 8b. But generally, for the comparison between two tables, it would be possible to conclude for which of two tables the departure from the EMH is greater if $\hat{\Gamma}^{(\lambda)}_{\rm EM}$ (for every λ) is always greater (or always less) for one table than for the other table.

Table 9: The values of $\hat{\Gamma}^{(\lambda)}_{\rm EM}$ applied to Tables 8a and 8b.

Values of λ	For Table 8a	For Table 8b
-0.5^*	0.040	0.042
0^*	0.066	0.068
0.5^*	0.081	0.082
1.0	0.089	0.088
1.5	0.091	0.090
2.0	0.089	0.088
2.5	0.084	0.084

$*$ indicates that $\hat{\Gamma}^{(\lambda)}_{\rm EM}$ is less for Table 8a than for Table 8b.

By the way, it is easily seen that the measure $\Gamma^{(0)}_{\rm EM}$ with equation (2.1) can be expressed as

$$(5.1)\qquad \Gamma^{(0)}_{\rm EM} = \frac{1}{2\log 2}\min_{\{D_i\}}\left[I^{(0)}\Big(\{G^*_{1(i)}\};\{D_i\}\Big) + I^{(0)}\Big(\{G^*_{2(i)}\};\{D_i\}\Big)\right],$$

where $\sum_{i=1}^{R-1} D_i = 1$ and $D_i > 0$. Therefore we point out that C_i in $\Gamma^{(\lambda)}_{\rm EM}$ is the value of D_i such that the sum of two KL distances (i.e., the KL distance between

$\{G^*_{1(i)}\}$ and $\{D_i\}$ with an EMH structure and the KL distance between $\{G^*_{2(i)}\}$ and $\{D_i\}$) is a minimum. [Note that the readers may also be interested in equation (5.1) with $I^{(0)}(\cdot;\cdot)$ replaced by the power-divergence $I^{(\lambda)}(\cdot;\cdot)$; however, it is difficult to obtain the value of D_i such that the corresponding two power-divergence is a minimum, and also difficult to obtain the maximum value of such a measure.]

Finally, we observe that

(i) the measure should be applied to contingency tables with *ordered* categories because it is not invariant under the same arbitrary permutations of row and column categories except the reverse order,

(ii) $\Gamma^{(\lambda)}_{\text{EM}}$ should be used when there is not a structure of EMH in square tables,

(iii) $\hat{\Gamma}^{(\lambda)}_{\text{EM}}$ cannot be used for testing goodness-of-fit of the EMH model (though $W^{(\lambda)}$ can be used),

and

(iv) the value of $\Gamma^{(1)}_{\text{EM}}$ is theoretically equal to the value of $\Gamma^{(2)}_{\text{EM}}$.

APPENDIX

The power-divergence statistic for testing goodness-of-fit of the EMH model is given by

$$W^{(\lambda)} = 2\,n\,I^{(\lambda)}\Big(\{\hat{p}_{ij}\};\{\hat{p}^{\text{ML}}_{ij}\}\Big) \qquad \text{for} \quad -\infty<\lambda<\infty \ ,$$

where

$$I^{(\lambda)}(\cdot;\cdot) = \frac{1}{\lambda(\lambda+1)} \sum_{i=1}^{R}\sum_{j=1}^{R} \hat{p}_{ij}\left[\left(\frac{\hat{p}_{ij}}{\hat{p}^{\text{ML}}_{ij}}\right)^{\lambda}-1\right], \qquad \hat{p}_{ij} = \frac{n_{ij}}{n}\ ,$$

and $\hat{p}^{\text{ML}}_{ij}$ is the maximum likelihood estimate of p_{ij} under the EMH model, where the values at $\lambda=-1$ and $\lambda=0$ are taken to be the limit as $\lambda\to-1$ and $\lambda\to 0$, respectively. The number of degrees of freedom is $R-2$.

ACKNOWLEDGMENTS

The authors thank a referee and the associate editor for their helpful comments.

REFERENCES

[1] Agresti, A. (1990). *Categorical Data Analysis*, Wiley, New York.

[2] Bhapkar, V.P. (1966). A note on the equivalence of two test criteria for hypotheses in categorical data, *Journal of the American Statistical Association*, **61**, 228–235.

[3] Bishop, Y.M.M.; Fienberg, S.E. and Holland, P.W. (1975). *Discrete Multivariate Analysis: Theory and Practice*, The MIT Press, Cambridge, Massachusetts.

[4] Cressie, N. and Read, T.R.C. (1984). Multinomial goodness-of-fit tests, *Journal of the Royal Statistical Society, Series B*, **46**, 440–464.

[5] Patil, G.P. and Taillie, C. (1982). Diversity as a concept and its measurement, *Journal of the American Statistical Association*, **77**, 548–561.

[6] Read, T.R.C. and Cressie, N. (1988). *Goodness-of-Fit Statistics for Discrete Multivariate Data*, Springer, New York.

[7] Stuart, A. (1955). A test for homogeneity of the marginal distributions in a two-way classification, *Biometrika*, **42**, 412–416.

[8] Tominaga, K. (1979). *Nippon no Kaisou Kouzou (Japanese Hierarchical Structure)*, University of Tokyo Press, Tokyo (in Japanese).

[9] Tomizawa, S. (1984). Three kinds of decompositions for the conditional symmetry model in a square contingency table, *Journal of the Japan Statistical Society*, **14**, 35–42.

[10] Tomizawa, S. (1995a). A generalization of the marginal homogeneity model for square contingency tables with ordered categories, *Journal of Educational and Behavioral Statistics*, **20**, 349–360.

[11] Tomizawa, S. (1995b). Measures of departure from marginal homogeneity for contingency tables with nominal categories, *Journal of the Royal Statistical Society, Series D; The Statistician*, **44**, 425–439.

[12] Tomizawa, S. and Makii, T. (2001). Generalized measures of departure from marginal homogeneity for contingency tables with nominal categories, *Journal of Statistical Research*, **35**, 1–24.

[13] Tomizawa, S.; Miyamoto, N. and Ashihara, N. (2003). Measure of departure from marginal homogeneity for square contingency tables having ordered categories, *Behaviormetrika*, **30**, 173–193.

REVSTAT – Statistical Journal
Volume 5, Number 3, November 2007, 285–304

A NOTE ON SECOND ORDER CONDITIONS IN EXTREME VALUE THEORY: LINKING GENERAL AND HEAVY TAIL CONDITIONS

Authors: M. Isabel Fraga Alves
– CEAUL, DEIO, Faculty of Science, University of Lisbon, Portugal
isabel.alves@fc.ul.pt

M. Ivette Gomes
– CEAUL, DEIO, Faculty of Science, University of Lisbon, Portugal
ivette.gomes@fc.ul.pt

Laurens de Haan
– Department of Economics, Erasmus University Rotterdam, The Netherlands
ldhaan@few.eur.nl

Cláudia Neves
– UIMA, Department of Mathematics, University of Aveiro, Portugal
claudia.neves@ua.pt

Received: July 2007 Revised: October 2007 Accepted: October 2007

Abstract:

- Second order conditions ruling the rate of convergence in any first order condition involving regular variation and assuring a unified extreme value limiting distribution function for the sequence of maximum values, linearly normalized, have appeared in several contexts whenever researchers are working either with a general tail, i.e., $\gamma \in \mathbb{R}$, or with heavy tails, with an extreme value index $\gamma > 0$. In this paper we shall clarify the link between the second order parameters, say ρ and $\widetilde{\rho}$ that have appeared in the two above mentioned set-ups, i.e., for a general tail and for heavy tails, respectively. We illustrate the theory with some examples and, for heavy tails, we provide a link with a third order framework.

Key-Words:

- *extreme value index; regular variation; semi-parametric estimation.*

AMS Subject Classification:

- Primary 62G32, 62E20, 26A12.

1. INTRODUCTION

Let $X_1, X_2, ..., X_n$ be an independent, identically distributed (i.i.d.) sample from an unknown distribution function (d.f.) F. It is well-known from Gnedenko's seminal work (Gnedenko, 1943) that if there exist normalizing constants $a_n > 0$, $b_n \in \mathbb{R}$ and a non-degenerate d.f. G such that, for all x,

$$\lim_{n\to\infty} P\Big\{a_n^{-1}\big(\max(X_1, ..., X_n) - b_n\big) \le x\Big\} = G(x) ,$$

G is, up to scale and location, an *Extreme Value* d.f., dependent on a shape parameter $\gamma \in \mathbb{R}$, and given by

$$G_\gamma(x) := \begin{cases} \exp\big(-(1+\gamma\,x)^{-1/\gamma}\big), & 1+\gamma\,x > 0 \quad \text{if } \gamma \neq 0 \\ \exp\big(-\exp(-x)\big), & x \in \mathbb{R} \quad \text{if } \gamma = 0 . \end{cases} \tag{1.1}$$

We then say that F is in the domain of attraction for maxima of the d.f. G_γ in (1.1) and write $F \in \mathcal{D}_{\mathcal{M}}(G_\gamma)$.

2. FIRST AND SECOND ORDER CONDITIONS

2.1. A general tail ($\gamma \in \mathbb{R}$)

The following *extended regular variation* property (de Haan, 1984), denoted ERV_γ, is a well-known necessary and sufficient condition for $F \in \mathcal{D}_{\mathcal{M}}(G_\gamma)$:

$$\lim_{t\to\infty} \frac{U(tx) - U(t)}{a(t)} = \begin{cases} \dfrac{x^\gamma - 1}{\gamma} & \text{if } \gamma \neq 0 \\ \ln x & \text{if } \gamma = 0 , \end{cases} \tag{2.1}$$

for every $x > 0$ and some positive measurable function a. For the case $\gamma > 0$ we see easily from (2.9) that we can choose $a(t) = \gamma\,U(t)$.

Apart from the first order condition in (2.1), we shall consider the most common second order condition, specifying the rate of convergence in (2.1). We shall assume the existence of a function $A(t)$, possibly not changing in sign and tending to zero as $t \to \infty$, such that

$$\lim_{t\to\infty} \frac{\frac{U(tx)-U(t)}{a(t)} - \frac{x^\gamma-1}{\gamma}}{A(t)} = H_{\gamma,\rho}(x) := \frac{1}{\rho}\left(\frac{x^{\gamma+\rho}-1}{\gamma+\rho} - \frac{x^\gamma - 1}{\gamma}\right) \tag{2.2}$$

for all $x > 0$, where $\rho \le 0$ is also a *second order* parameter controlling the speed of convergence of maximum values, linearly normalized, towards the limit law in

(1.1), for a general $\gamma \in \mathbb{R}$. We then say that the function U is of *second order extended regular variation, and use the notation* $U \in 2ERV_{\gamma,\rho}$. In (2.2), the cases $\gamma = 0$ and $\rho = 0$ are obtained by continuity arguments. More specifically, we can write

$$H_{\gamma,\rho}(x) = \begin{cases} \dfrac{1}{\rho}\left(\dfrac{x^\rho - 1}{\rho} - \ln x\right) & \text{if } \gamma = 0,\ \rho \neq 0 \\ \dfrac{1}{\gamma}\left(x^\gamma \ln x - \dfrac{x^\gamma - 1}{\gamma}\right) & \text{if } \gamma \neq 0,\ \rho = 0 \\ \dfrac{\ln^2 x}{2} & \text{if } \gamma = \rho = 0 . \end{cases}$$

We remark that $|A| \in RV_\rho$. For a large variety of models we have $\rho < 0$ thus making sensible to simplify (2.2). We now state:

Proposition 2.1 (Gomes and Neves, 2007)**.** *Let us assume that there exist* $a(\cdot)$ *and* $A(\cdot)$ *such that* (2.2) *holds, with* $\rho < 0$*. Then, there exist* $a_0(\cdot)$ *and* $A_0(\cdot)$ *such that*

$$\lim_{t\to\infty} \frac{\frac{U(tx)-U(t)}{a_0(t)} - \frac{x^\gamma - 1}{\gamma}}{A_0(t)} = \frac{x^{\gamma+\rho} - 1}{\gamma+\rho} \tag{2.3}$$

with

$$A_0(t) = A(t)/\rho , \qquad a_0(t) = a(t)\big(1 - A_0(t)\big) . \tag{2.4}$$

From Theorem A in Draisma de Haan, Peng and Pereira (1999), with slight additions in Ferreira, de Haan and Peng (2003) and in de Haan and Ferreira (2006), we state the following:

Theorem 2.1. *Suppose the right endpoint* $x^F := U(\infty) > 0$ *and there exist* $a(\cdot)$ *and* $A(\cdot)$ *such that* (2.2) *holds, with* $\rho \le 0$, $\gamma \neq \rho$*. Define*

$$\overline{A}(t) := \left(\frac{a(t)}{U(t)} - \gamma_+\right), \qquad \gamma_+ := \max(0,\gamma) . \tag{2.5}$$

Then for $\gamma + \rho < 0$

$$l := \lim_{t\to\infty}\left(U(t) - \frac{a(t)}{\gamma}\right) \quad \textit{exists and is finite} \tag{2.6}$$

and the following holds

$$\overline{A}(t) \underset{t\to\infty}{\longrightarrow} 0 \qquad \textit{and} \qquad \frac{\overline{A}(t)}{A(t)} \underset{t\to\infty}{\longrightarrow} c ,$$

with

$$c = \begin{cases} 0 & \textit{if } \ \gamma < \rho \le 0 \\ \dfrac{\gamma}{\gamma+\rho} & \textit{if } \ 0 \le -\rho < \gamma \ \textit{ or } \ (0 < \gamma < -\rho \ \textit{ and } \ l = 0) \\ \pm\infty & \textit{if } \ \gamma+\rho = 0 \ \textit{ or } \ (0 < \gamma < -\rho \ \textit{ and } \ l \neq 0) \ \textit{ or } \ \rho < \gamma \le 0 . \end{cases} \tag{2.7}$$

2.1.1. Heavy tails $(\gamma > 0)$

The most typical first order condition for heavy tails, i.e., for the case $\gamma > 0$ in (1.1), comes also from Gnedenko (1943). For any real τ, let us denote by RV_τ the class of regularly varying functions with an index of regular variation τ, i.e., positive measurable functions g such that $\lim_{t\to\infty} g(tx)/g(t) = x^\tau$ for all $x > 0$. Then, for $\gamma > 0$,

$$F \in \mathcal{D}_{\mathcal{M}}(G_\gamma) \iff \overline{F} = 1-F \in RV_{-1/\gamma} \; . \tag{2.8}$$

Equivalently, and with U standing for a quantile type function associated to F and defined by $U(t) := \big(1/(1-F)\big)^{\leftarrow}(t) = \inf\big\{x\colon F(x) \geq 1-\frac{1}{t}\big\}$, de Haan (1970) established that

$$F \in \mathcal{D}_{\mathcal{M}}(G_\gamma) \iff U \in RV_\gamma \; . \tag{2.9}$$

To measure the rate of convergence in (2.9), it is then sensible to consider one of the following conditions:

$$\begin{aligned} &\lim_{t\to\infty} \frac{\frac{U(tx)}{U(t)} - x^\gamma}{\widetilde{A}(t)} = x^\gamma \, \frac{x^{\widetilde{\rho}}-1}{\widetilde{\rho}} \iff \\ &\iff \lim_{t\to\infty} \frac{\ln U(tx) - \ln U(t) - \gamma \ln x}{\widetilde{A}(t)} = \frac{x^{\widetilde{\rho}}-1}{\widetilde{\rho}} \; , \end{aligned} \tag{2.10}$$

for all $x > 0$, where $\widetilde{\rho} \leq 0$ is a *second order* parameter controlling the speed of convergence of maximum values, linearly normalized, towards the limit law in (1.1) pertaining to $\gamma > 0$. Under these circumstances, we say that the function U is of *regular variation of second order*, and use the notation $U \in 2RV(\gamma, \widetilde{\rho})$. We remark that $|\widetilde{A}| \in RV_{\widetilde{\rho}}$.

3. THE LINK BETWEEN THE SECOND ORDER CONDITION FOR A HEAVY AND FOR A GENERAL TAIL

The following results hold with any measurable (eventually) positive function U.

Lemma 3.1. *If* (2.1) *holds for some* $\gamma \in \mathbb{R}$, *then the auxiliary function* $a(t)$ *in* (2.1) *is of regular variation at infinity with index* γ, *i.e.,* $a \in RV_\gamma$ *and*

$$\lim_{t\to\infty} \frac{a(t)}{U(t)} = \gamma_+ := \max(0,\gamma) \; .$$

Moreover, if $\gamma > 0$, both functions a and U belong to RV_γ; if $\gamma < 0$, then $x^F = U(\infty) < \infty$, $\lim_{t\to\infty} a(t)/\big(x^F - U(t)\big) = -\gamma$ and $x^F - U \in RV_\gamma$.

Furthermore, with $\gamma_- := \min(\gamma, 0)$, and provided that $U(\infty) > 0$,

$$\lim_{t\to\infty} \frac{\ln U(tx) - \ln U(t)}{a(t)/U(t)} = \frac{x^{\gamma_-} - 1}{\gamma_-}, \qquad \textit{for every } x > 0\ . \tag{3.1}$$

Proof: The first part of the lemma comes from Theorems 1.9 and 1.10 in Geluk and de Haan (1987). The limit in (3.1) follows easily when we distinguish between the cases $\gamma > 0$ and $\gamma \le 0$. □

For the derivation of asymptotic properties of semi-parametric estimators of γ, a topic out of the scope of this paper, it is important to know, for all $x > 0$, not only the rate of convergence of $\ln U(tx) - \ln U(t)$, but also of $U(tx)/U(t)$ and of $U(t)/U(tx)$, as $t \to \infty$. We shall now see in more detail for the different relevant subspaces of the semi-plane $(\gamma, \rho) \in \mathbb{R}\times\mathbb{R}_0^-$, the limiting behaviour, as $t \to \infty$, of $U(tx)/U(t)$ and $U(t)/U(tx)$. The limit behavior of $\ln U(tx) - \ln U(t)$ has been analyzed e.g. in Appendix B of de Haan and Ferreira (2006).

Lemma 3.2. *Assume that (2.2) holds, i.e., $U \in 2ERV_{\gamma,\rho}$. Then, we may write*

$$\frac{U(tx)}{U(t)} = x^{\gamma_+} + \overline{A}(t)\left[\frac{x^\gamma - 1}{\gamma} + A(t)\,\overline{a}(x,t;\gamma,\rho)\,\big(1 + o(1)\big)\right], \tag{3.2}$$

where

$$\overline{a}(x,t;\gamma,\rho) = \begin{cases} \dfrac{\ln^2 x}{2} & \textit{if } \gamma = \rho = 0 \\[2mm] \dfrac{1}{\gamma}\left(x^\gamma \ln x - \dfrac{x^\gamma - 1}{\gamma}\right) & \textit{if } \gamma < \rho = 0 \\[2mm] \dfrac{1}{\rho}\left(\dfrac{x^{\gamma+\rho} - 1}{\gamma+\rho} - \dfrac{x^\gamma - 1}{\gamma}\right) & \textit{if } \gamma \le 0,\ \rho < 0 \\[2mm] \dfrac{\gamma}{\rho\,\overline{A}(t)}\left(\dfrac{x^{\gamma+\rho} - 1}{\gamma+\rho} - \dfrac{x^\gamma - 1}{\gamma}\right) & \textit{if } \gamma > 0,\ \rho < 0 \\[2mm] \dfrac{1}{\overline{A}(t)}\left(x^\gamma \ln x - \dfrac{x^\gamma - 1}{\gamma}\right) & \textit{if } \rho = 0 < \gamma\ . \end{cases}$$

Proof: Directly from (2.2), we get

$$\frac{U(tx)}{U(t)} - 1 = \frac{a(t)}{U(t)}\left\{\frac{x^\gamma - 1}{\gamma} + \frac{A(t)}{\rho}\left(\frac{x^{\gamma+\rho} - 1}{\gamma+\rho} - \frac{x^\gamma - 1}{\gamma}\right)\big(1 + o(1)\big)\right\}.$$

With the notation in (2.5), i.e., $a(t)/U(t) = \gamma_+ + \overline{A}(t)$, we may write

$$\frac{U(tx)}{U(t)} - 1 \;=\; \gamma_+\left(\frac{x^\gamma-1}{\gamma}\right) + \overline{A}(t)\left(\frac{x^\gamma-1}{\gamma}\right) + \frac{A(t)}{\rho}\left(\frac{x^{\gamma+\rho}-1}{\gamma+\rho} - \frac{x^\gamma-1}{\gamma}\right)\big(\gamma_+ + \overline{A}(t)\big)\big(1+o(1)\big)$$

and (3.2) follows for any $\rho < 0$.

If $\rho = 0$ and $\gamma \neq 0$, then, also directly from (2.2), and by continuity arguments,

$$\frac{U(tx)}{U(t)} - 1 \;=\; \frac{a(t)}{U(t)}\left\{\frac{x^\gamma-1}{\gamma} + \frac{A(t)}{\gamma}\left(x^\gamma \ln x - \frac{x^\gamma-1}{\gamma}\right)\big(1+o(1)\big)\right\},$$

and things work as before, with $A(t)/\rho$ replaced by $A(t)/\gamma$ and $\frac{x^{\gamma+\rho}-1}{\gamma+\rho}$ replaced by $x^\gamma \ln x$. The case $\gamma = \rho = 0$ comes again directly from (2.2) and by continuity arguments. □

Theorem 3.1. *Let $U \in ERV_{\gamma,\rho}$ as introduced in* (2.2). *Let c be the limit in* (2.7).

(i) *If $\gamma > 0$,*

$$\lim_{t\to\infty} \frac{\frac{U(t)}{U(tx)} - x^{-\gamma}}{\widetilde{A}(t)} \;=\; K_{\gamma,\rho}(x) \;:=\; \begin{cases} -x^{-\gamma}\,\dfrac{x^\rho-1}{\rho} & \text{if } \; c = \dfrac{\gamma}{\gamma+\rho} \\[2ex] -x^{-\gamma}\,\dfrac{x^{-\gamma}-1}{-\gamma} & \text{if } \; c = \pm\infty\,, \end{cases} \tag{3.3}$$

for all $x > 0$, where, with $\overline{A}(t)$ given in (2.5),

$$\widetilde{A}(t) \;:=\; \begin{cases} \dfrac{\gamma\,A(t)}{\gamma+\rho} & \text{if } \; c = \dfrac{\gamma}{\gamma+\rho} \\[2ex] \overline{A}(t) & \text{if } \; c = \pm\infty\,. \end{cases} \tag{3.4}$$

Necessarily, $|\widetilde{A}| \in RV_{\widetilde{\rho}}$, with

$$\widetilde{\rho} \;=\; \begin{cases} \rho & \text{if } \; c = \dfrac{\gamma}{\gamma+\rho} \\[2ex] -\gamma & \text{if } \; c = \pm\infty\,. \end{cases} \tag{3.5}$$

(ii) *If $\gamma \le 0$, we need further to assume that $\gamma \neq \rho$. Then,*

$$\lim_{t\to\infty} \frac{\frac{U(t)}{a^*(t)}\left(1 - \frac{U(t)}{U(tx)}\right) - \frac{x^\gamma-1}{\gamma}}{A^*(t)} = K^*_{\gamma,\rho}(x) \;=\; \begin{cases} x^\gamma \ln x & \text{if } \; \gamma<\rho=0 \\[1ex] \dfrac{x^{\gamma+\rho}-1}{\gamma+\rho} & \text{if } \; \gamma<\rho<0 \\[2ex] \dfrac{x^{2\gamma}-1}{2\,\gamma} & \text{if } \; \rho<\gamma<0 \\[2ex] \ln^2 x & \text{if } \; \rho<\gamma=0\,, \end{cases} \tag{3.6}$$

where

$$A^*(t) \;=\; \begin{cases} \dfrac{A(t)}{\gamma} & \text{if } \;\gamma<\rho=0 \\ \dfrac{A(t)}{\rho} & \text{if } \;\gamma<\rho<0 \\ -\dfrac{2\,\overline{A}(t)}{\gamma} & \text{if } \;\rho<\gamma<0 \\ -\overline{A}(t) & \text{if } \;\rho<\gamma=0 \end{cases} \tag{3.7}$$

and

$$a^*(t) \;=\; \begin{cases} a(t) & \text{if } \;\rho<\gamma=0 \\ a(t)\,\bigl(1-A^*(t)\bigr) & \text{otherwise .} \end{cases} \tag{3.8}$$

Necessarily, $|A^*| \in RV_{\rho^*}$*, with*

$$\rho^* \;=\; \begin{cases} \rho & \text{if } \;\gamma<\rho\le 0 \\ \gamma & \text{if } \;\rho<\gamma\le 0 \;. \end{cases} \tag{3.9}$$

Proof: We shall consider the cases (i) and (ii) separately.

Case (i): If $\gamma>0$, $\rho<0$ and (2.2) holds, i.e., $U\in 2ERV_{\gamma,\rho}$, we have from (3.2),

$$\frac{U(tx)}{U(t)} - x^\gamma \;=\; \overline{A}(t)\left(\frac{x^\gamma-1}{\gamma}\right) + \frac{\gamma\,A(t)}{\rho}\left(\frac{x^{\gamma+\rho}-1}{\gamma+\rho} - \frac{x^\gamma-1}{\gamma}\right) + o\bigl(A(t)\bigr)\;.$$

If $c=\pm\infty$, then $A(t)=o\bigl(\overline{A}(t)\bigr)$,

$$\frac{U(tx)}{U(t)} - x^\gamma \;=\; x^\gamma\left(\frac{x^{-\gamma}-1}{-\gamma}\right)\overline{A}(t) + o\bigl(\overline{A}(t)\bigr) \quad\text{and}\quad \frac{\frac{U(tx)}{U(t)}-x^\gamma}{\overline{A}(t)} \underset{t\to\infty}{\longrightarrow} x^\gamma\left(\frac{x^{-\gamma}-1}{-\gamma}\right)\;.$$

If $c=\gamma/(\gamma+\rho)$, we get $\overline{A}(t)=\frac{\gamma\,A(t)}{\gamma+\rho}\bigl(1+o(1)\bigr)$. Since in this region $\gamma\neq-\rho$, we may further write

$$\begin{aligned} \frac{U(tx)}{U(t)} - x^\gamma \;&=\; x^\gamma\left(\overline{A}(t)\left(\frac{x^{-\gamma}-1}{-\gamma}\right) + \frac{\gamma\,A(t)}{\gamma+\rho}\left(\frac{x^\rho-1}{\rho} - \frac{x^{-\gamma}-1}{-\gamma}\right) + o\bigl(A(t)\bigr)\right) \\ &=\; \frac{\gamma\,A(t)}{\gamma+\rho}\,x^\gamma\left(\frac{x^\rho-1}{\rho}\right) + o\bigl(A(t)\bigr)\;. \end{aligned}$$

Consequently,

$$\lim_{t\to\infty} \frac{\frac{U(tx)}{U(t)}-x^\gamma}{\widetilde{A}(t)} \;=\; \begin{cases} x^\gamma\,\dfrac{x^\rho-1}{\rho} & \text{if } \;c=\dfrac{\gamma}{\gamma+\rho} \\ x^\gamma\,\dfrac{x^{-\gamma}-1}{-\gamma} & \text{if } \;c=\pm\infty \end{cases} \;=\; -x^{2\gamma}\,K_{\gamma,\rho}(x)\;, \tag{3.10}$$

with $K_{\gamma,\rho}(x)$ and $\widetilde{A}(t)$ defined in (3.3) and (3.4), respectively. Finally, (3.10), together with the fact that

$$\frac{U(tx)}{U(t)} - x^\gamma \;=\; -x^\gamma \,\frac{U(tx)}{U(t)} \left(\frac{U(t)}{U(tx)} - x^{-\gamma}\right) \;=\; -x^{2\gamma}\left(\frac{U(t)}{U(tx)} - x^{-\gamma}\right)\big(1+o(1)\big)\;,$$

leads us to the limit in (3.3), with $\widetilde{A}(t)$ and $\widetilde{\rho}$ given in (3.4) and (3.5), respectively.

If $\gamma > 0$ and $\rho = 0$, we get, again from (3.2),

$$\begin{aligned}\frac{U(tx)}{U(t)} - x^\gamma \;&=\; \overline{A}(t)\left(\frac{x^\gamma-1}{\gamma}\right) + A(t)\left(x^\gamma \ln x - \frac{x^\gamma-1}{\gamma}\right) + o\big(A(t)\big)\\ &=\; x^\gamma\left(\overline{A}(t)\left(\frac{x^{-\gamma}-1}{-\gamma}\right) + A(t)\left(\ln x - \frac{x^{-\gamma}-1}{-\gamma}\right) + o\big(A(t)\big)\right).\end{aligned}$$

But if $\gamma > 0$ and $\rho = 0$, then $c = \gamma/(\gamma+\rho) = 1$, $\overline{A}(t) = A(t) + o\big(A(t)\big)$, and

$$\frac{U(tx)}{U(t)} - x^\gamma \;=\; A(t)\, x^\gamma \ln x + o\big(A(t)\big)\;.$$

Consequently, (3.3) holds, with $\widetilde{A}(t) = A(t) \equiv \gamma\, A(t)/(\gamma+\rho)$ and $\widetilde{\rho} = \rho = 0$.

Case (ii): If $\gamma < \rho = 0$, we get, again from (3.2),

$$\begin{aligned}\frac{U(t)}{a(t)}\left(1 - \frac{U(t)}{U(tx)}\right) \;&=\; \frac{x^\gamma-1}{\gamma} + \frac{A(t)}{\gamma}\left(x^\gamma \ln x - \frac{x^\gamma-1}{\gamma}\right) + o\big(A(t)\big)\\ &=\; \frac{x^\gamma-1}{\gamma}\left(1 - \frac{A(t)}{\gamma}\right) + \frac{A(t)}{\gamma}\, x^\gamma \ln x + o\big(A(t)\big)\end{aligned}$$

and with $a^*(t) = a(t)\left(1 - \frac{A(t)}{\gamma}\right)$,

$$\frac{U(t)}{a^*(t)}\left(1 - \frac{U(t)}{U(tx)}\right) = \frac{x^\gamma-1}{\gamma} + \frac{A(t)}{\gamma}\, x^\gamma \ln x + o\big(A(t)\big)\;.$$

Consequently, (3.6), (3.7), (3.8) and (3.9) follow in this region of the (γ,ρ)-plane.

If $\gamma < \rho < 0$, $a(t)/U(t) \equiv \overline{A}(t) = o(A(t))$, and again from (3.2),

$$\begin{aligned}\frac{U(t)}{a(t)}\left(1 - \frac{U(t)}{U(tx)}\right) \;&=\; \frac{x^\gamma-1}{\gamma} + \frac{A(t)}{\rho}\left(\frac{x^{\gamma+\rho}-1}{\gamma+\rho} - \frac{x^\gamma-1}{\gamma}\right) + o\big(A(t)\big)\\ &=\; \frac{x^\gamma-1}{\gamma}\left(1 - \frac{A(t)}{\rho}\right) + \frac{A(t)}{\rho}\left(\frac{x^{\gamma+\rho}-1}{\gamma+\rho}\right) + o\big(A(t)\big)\end{aligned}$$

and with $a^*(t) = a(t)\left(1-\frac{A(t)}{\rho}\right)$,

$$\frac{U(t)}{a^*(t)}\left(1-\frac{U(t)}{U(tx)}\right) = \frac{x^\gamma-1}{\gamma} + \frac{A(t)}{\rho}\left(\frac{x^{\gamma+\rho}-1}{\gamma+\rho}\right) + o\big(A(t)\big)~,$$

and the results in the proposition hold.

If $\rho<\gamma\le 0$, $A(t)=o\left(\overline{A}(t)\right)$, and also from (3.2), we get

$$\frac{U(t)}{U(tx)} = 1-\overline{A}(t)\left(\frac{x^\gamma-1}{\gamma}\right) + \overline{A}^2(t)\left(\frac{x^\gamma-1}{\gamma}\right)^2\big(1+o(1)\big)~. \tag{3.11}$$

Consequently, for $\gamma<0$, since $\left(\frac{x^\gamma-1}{\gamma}\right)^2 = \frac{2}{\gamma}\left(\frac{x^{2\gamma}-1}{2\gamma}-\frac{x^\gamma-1}{\gamma}\right)$

$$\begin{aligned}\frac{U(t)}{a(t)}\left(1-\frac{U(t)}{U(tx)}\right) &= \frac{x^\gamma-1}{\gamma} - \frac{2\,\overline{A}(t)}{\gamma}\left(\frac{x^{2\gamma}-1}{2\,\gamma}-\frac{x^\gamma-1}{\gamma}\right)\big(1+o(1)\big)\\ &= \frac{x^\gamma-1}{\gamma}\left(1+\frac{2\,\overline{A}(t)}{\gamma}\right) - \frac{2\,\overline{A}(t)}{\gamma}\left(\frac{x^{2\gamma}-1}{2\,\gamma}\right)\big(1+o(1)\big)~,\end{aligned}$$

and with $a^*(t) = a(t)\left(1+\frac{2\,\overline{A}(t)}{\gamma}\right)$,

$$\frac{U(t)}{a^*(t)}\left(1-\frac{U(t)}{U(tx)}\right) = \frac{x^\gamma-1}{\gamma} - \frac{2\,\overline{A}(t)}{\gamma}\left(\frac{x^{2\gamma}-1}{2\,\gamma}\right)\big(1+o(1)\big)~,$$

and the results in the proposition follow.

If $\rho<\gamma=0$, then from (3.11), we get

$$\frac{U(t)}{a(t)}\left(1-\frac{U(t)}{U(tx)}\right) = \ln x - \overline{A}(t)\,\ln^2 x\,\big(1+o(1)\big)~,$$

and the result in the proposition follows as well. □

Corollary 3.1. *Under the conditions and notations of Proposition 2.1, and for* $\gamma>0$,

$$\lim_{t\to\infty}\frac{\ln U(tx)-\ln U(t)-\gamma\ln x}{\widetilde{A}(t)} = \widetilde{K}_{\gamma,\rho}(x) := \begin{cases}\dfrac{x^\rho-1}{\rho} & \textit{if } c=\dfrac{\gamma}{\gamma+\rho}\\[2ex] \dfrac{x^{-\gamma}-1}{-\gamma} & \textit{if } c=\pm\infty~,\end{cases} \tag{3.12}$$

for every $x>0$, *and with* $\widetilde{A}$ *provided in* (3.4).

Proof: The proof follows immediately from relation (3.3). □

Remark 3.1. Note that the second order condition in (3.12) is the usual second order condition for heavy tails, i.e., the second order condition provided in (2.10).

Remark 3.2. Note next that the region $\{(\gamma,\rho)\colon 0<\gamma<-\rho \text{ and } l\neq 0\}$ in the (γ,ρ)-plane, jointly with the line $\rho=-\gamma$, are transformed in the line $\widetilde{\rho}=-\gamma$ in the $(\gamma,\widetilde{\rho})$-plane. There we have $c=\pm\infty$. Outside that line we have $c=\gamma/(\gamma+\widetilde{\rho})=\gamma/(\gamma+\rho)$.

Remark 3.3. For $\gamma>0$, the rate of convergence in (3.1), i.e., the rate of convergence of $\big(\ln U(tx)-\ln U(t)\big)/\big(a(t)/U(t)\big)-\ln x$ towards zero, is measured by $\widetilde{A}(t)$ in (3.4) only if $\rho\neq 0$. If $\rho=0$, the rate of convergence in (3.1) can be of a smaller order than $\widetilde{A}(t)$ as may be seen in Example 4.1.
For $\gamma\leq 0$, Lemma 3.2 gives rise to (3.1) in a similar way as it yields Corollary 3.1.

4. EXAMPLES AND SOME ADDITIONAL COMMENTS

Example 4.1. (A model with $\rho=\widetilde{\rho}=0$ and $\gamma>0$). Consider the model $U(t)=t^\gamma(1+\ln t)$. Then

$$U(tx)-U(t) \;=\; \gamma\, t^\gamma(\ln t+1)\left(\frac{x^\gamma-1}{\gamma}+\frac{x^\gamma\ln x}{\gamma(\ln t+1)}\right)\,, \qquad x>0\,,$$

i.e., $\rho=0$ in (2.2), since

$$\frac{\frac{U(tx)-U(t)}{\gamma\, t^\gamma(\ln t+1)}-\frac{x^\gamma-1}{\gamma}}{1/\big(\gamma(\ln t+1)\big)} \;=\; x^\gamma\ln x\,.$$

Notice that $H_{\gamma,0}(x)=\gamma^{-1}\big(x^\gamma\ln x-(x^\gamma-1)/\gamma\big)$, meaning that (2.2) is equivalent to

$$\frac{\frac{U(tx)-U(t)}{a(t)\,(1-A(t)/\gamma)}-\frac{x^\gamma-1}{\gamma}}{A(t)/\gamma} \;\underset{t\to\infty}{\longrightarrow}\; x^\gamma\ln x\,,$$

as stated in (2.3). Consequently we should choose

$$A(t)=\frac{1}{\ln t+1}\in RV_0\,, \qquad a(t)\left(1-\frac{1}{\gamma(\ln t+1)}\right) \;=\; \gamma\, t^\gamma(\ln t+1)\,.$$

Theorem 2.1 yields $c=1$ while Theorem 3.1 determines $\widetilde{\rho}=0$ and $\widetilde{A}(t)=A(t)$. Indeed, we have

$$\text{(4.1)}\qquad \ln U(tx)-\ln U(t)-\gamma\ln x \;=\; \frac{\ln x}{\ln t+1}+\frac{\ln^2 x}{2(\ln t+1)^2}+o\left(\frac{1}{\ln^2 t}\right),$$

as $t \to \infty$, thus making suitable to take $A(t) = (\ln t + 1)^{-1}$ in the left hand side of

$$\lim_{t\to\infty} \frac{\ln U(tx) - \ln U(t) - \gamma \ln x}{A(t)} = \ln x \ , \qquad x > 0$$

and (3.12) holds in fact with $\widetilde{A}(t) = A(t)$. Furthermore, after a few manipulations of (4.1), we get

$$\frac{\frac{\ln U(tx) - \ln U(t)}{\gamma\left(1+\frac{1}{\gamma(\ln t+1)}\right)} - \ln x}{\frac{1}{2\ln^2 t}} \underset{t\to\infty}{\longrightarrow} \ln^2 x \ .$$

Therefore, the rate of convergence in (3.1) is of the order of $1/\ln^2 t = o\big(A(t)\big)$, as mentioned in Remark 3.3.

Example 4.2. For the Fréchet model, $F(x) = \exp\left(-x^{-1/\gamma}\right)$, $x \geq 0$ $(\gamma > 0)$, we get successively,

$$\begin{aligned} U(t) &= \left(-\ln\left(1 - \frac{1}{t}\right)\right)^{-\gamma} \\ &= t^\gamma \left(1 + \frac{1}{2t} + \frac{1}{3t^2} + o\big(t^{-2}\big)\right)^{-\gamma} \\ &= t^\gamma \left(1 - \frac{\gamma}{2t} + \frac{\gamma(3\gamma - 5)}{24t^2} + o\big(t^{-2}\big)\right) . \end{aligned}$$

Hence,

$$U(tx) - U(t) = \begin{cases} \gamma\, t^\gamma \left(\dfrac{x^\gamma - 1}{\gamma} - \dfrac{\gamma - 1}{2t}\left(\dfrac{x^{\gamma-1} - 1}{\gamma - 1}\right) + o\big(t^{-1}\big)\right) & \text{if } \gamma \neq 1 \\ t\left((x-1) - \dfrac{1}{12\,t^2}\,(x^{-1} - 1) + o\big(t^{-2}\big)\right) & \text{if } \gamma = 1 \ . \end{cases}$$

If we make correspondence with condition (2.3), we see that $\rho = \begin{cases} -1 & \text{if } \gamma \neq 1 \\ -2 & \text{if } \gamma = 1 \ . \end{cases}$ Likewise, (2.4) can be set as

$$a_0(t) = \gamma\, t^\gamma \qquad \text{and} \qquad A_0(t) = \begin{cases} \dfrac{1-\gamma}{2t} & \text{if } \gamma \neq 1 \\ \dfrac{1}{12\,t^2} & \text{if } \gamma = 1 \ . \end{cases}$$

According to Proposition 2.1, if we choose

$$A(t) = \rho A_0(t) = \begin{cases} \dfrac{\gamma - 1}{2t} & \text{if } \gamma \neq 1 \\ -\dfrac{1}{6\,t^2} & \text{if } \gamma = 1 \end{cases}$$

and

$$a(t) \;=\; \gamma\, t^\gamma/\big(1 - A_0(t)\big) \;=\; \begin{cases} \dfrac{2\,\gamma\, t^{\gamma+1}}{2\,t+\gamma-1} & \text{if } \gamma \neq 1 \\[2ex] \dfrac{12\,t^3}{12\,t^2-1} & \text{if } \gamma = 1\ , \end{cases}$$

we get the limiting result in (2.2).

We will derive that (3.12) holds for $\widetilde{A}(t) = \gamma/(2\,t)$, with $\widetilde{\rho} = -1 \neq \rho = -2$ for $\gamma = 1$, and $\widetilde{\rho} = -1 = \rho$ for $\gamma \neq 1$. As seen before regarding the limit in (2.2), we have whenever $\gamma \neq 1$,

$$\begin{aligned} U(t) \;&=\; t^\gamma \left(1 - \frac{\gamma}{2\,t} + \frac{\gamma\,(3\,\gamma-5)}{24\,t^2} + o\big(t^{-2}\big)\right)\ ; \\ a(t) \;&=\; 2\,\gamma\, t^{\gamma+1}/(2\,t+\gamma+\rho)\ ; \\ A(t) \;&=\; -\rho\,(\gamma+\rho)/(2\,t)\ . \end{aligned}$$

Then,

$$\begin{aligned} \overline{A}(t) \;=\; \frac{a(t)}{U(t)} - \gamma \;&=\; \frac{2\,\gamma\, t}{2\,t+\gamma+\rho} \left(1 + \frac{\gamma}{2\,t} - \frac{\gamma\,(3\,\gamma-5)}{24\,t^2} + \frac{\gamma^2}{4\,t^2} + o\big(t^{-2}\big)\right) - \gamma \\ &=\; \frac{2\,\gamma\, t\,(2\,t+\gamma)}{2\,t\,(2\,t+\gamma+\rho)} - \gamma - \frac{2\,\gamma^2\, t\,(9\,\gamma-5)}{24\,t^2\,(2\,t+\gamma+\rho)} + o\big(t^{-2}\big) \\ &=\; -\frac{\gamma\,\rho}{2\,t+\gamma+\rho} \left(1 + \frac{\gamma\,(9\,\gamma-5)}{12\,\rho\, t} + o\big(t^{-1}\big)\right) \underset{t\to\infty}{\longrightarrow} 0\ , \end{aligned}$$

and

$$\frac{\overline{A}(t)}{A(t)} \;=\; \frac{2\,\gamma\, t}{(\gamma+\rho)\,(2\,t+\gamma+\rho)} \left(1 + \frac{\gamma\,(9\,\gamma-5)}{12\,t} + o\big(t^{-1}\big)\right) \underset{t\to\infty}{\longrightarrow} \frac{\gamma}{\gamma+\rho}\ .$$

Let us think on

$$\begin{aligned} U(t) - \frac{a(t)}{\gamma} \;&=\; -\frac{U(t)}{\gamma}\,\overline{A}(t) \\ &=\; t^\gamma \left(\frac{2\,\rho\, t - \gamma(\gamma+\rho)}{2\,t\,(2\,t+\gamma+\rho)} + \frac{\gamma\,(3\gamma-5)}{24\,t^2} + o\big(t^{-2}\big)\right) \underset{t\to\infty}{\longrightarrow} 0 =: l, \quad \text{if } 0<\gamma<1\,. \end{aligned}$$

Hence, we conclude that $c = \gamma/(\gamma+\rho)$, for $\gamma \neq 1$.

If we consider the case $\gamma = 1$,

$$\begin{aligned} \frac{a(t)}{U(t)} \;&=\; \frac{12\,t^2}{12\,t^2-1} \left(1 + \frac{1}{2\,t} + \frac{1}{12\,t^2} + \frac{1}{4\,t^2} + o\big(t^{-2}\big)\right) \\ &=\; 1 + \frac{1}{2\,t} + \frac{5}{12\,t^2} + o\big(t^{-2}\big)\ . \end{aligned}$$

Consequently, and as was expected from Theorem 2.1,

$$\overline{A}(t) \;=\; \frac{a(t)}{U(t)} - 1 \;=\; \frac{1}{2\,t}\left(1 + \frac{5}{6\,t} + o\big(t^{-1}\big)\right) \underset{t\to\infty}{\longrightarrow} 0 ,$$

$$\frac{\overline{A}(t)}{A(t)} \;=\; -3\,t\left(1 + \frac{5}{6\,t} + o\big(t^{-1}\big)\right) \underset{t\to\infty}{\longrightarrow} -\infty, \qquad \text{i.e.,}\;\; c = -\infty$$

and

$$U(t) - a(t) \;=\; -\frac{1}{2} - \frac{t}{12\,t^2 - 1} - \frac{1}{12\,t} + o\big(t^{-1}\big) \underset{t\to\infty}{\longrightarrow} -\frac{1}{2} = l .$$

Since this limit l is different from zero and $\gamma = 1 < -\rho = 2$, we indeed expected to have $c = \pm\infty$, as actually happens. Now, from Theorem 3.1, $\widetilde{\rho} = -\gamma = -1$ and we may choose

$$\widetilde{A}(t) \;=\; \overline{A}(t) \;=\; \frac{1}{2\,t}\left(1 + \frac{5}{6\,t} + o\big(t^{-1}\big)\right) ,$$

or more simply $\widetilde{A}(t) = 1/(2\,t)$. Indeed, and as mentioned before for $\gamma = 1$, (3.12) holds true with $\widetilde{A}(t) = \gamma/(2\,t)$ and $\widetilde{\rho} = -1 \neq \rho = -2$.

Example 4.3. Consider the extreme value model with d.f. $G_\gamma(x) = \exp\big(-(1+\gamma\,x)^{-1/\gamma}\big)$, $1+\gamma\,x > 0$, $\gamma \in \mathbb{R}$. For this model,

$$\begin{aligned}
U(t) \;&=\; \frac{\Big(-\ln\big(1-\frac{1}{t}\big)\Big)^{-\gamma} - 1}{\gamma} \\
&=\; \frac{t^\gamma}{\gamma}\left(1 - t^{-\gamma} - \frac{\gamma}{2\,t} + \frac{\gamma\,(3\,\gamma-5)}{24\,t^2} + o\big(t^{-2}\big)\right) \\
&=\; \begin{cases}
-\dfrac{1}{\gamma}\left(1 - t^\gamma + \dfrac{\gamma\,t^{\gamma-1}}{2} + o\big(t^{\gamma-1}\big)\right) & \text{if } \gamma < 0 \\[2ex]
\ln t - \dfrac{1}{2\,t} + o\big(t^{-1}\big) & \text{if } \gamma = 0 \\[2ex]
\dfrac{t^\gamma}{\gamma}\left(1 - t^{-\gamma} - \dfrac{\gamma}{2\,t} + o\big(t^{-1}\big)\right) & \text{if } 0 < \gamma < 1 \\[2ex]
\dfrac{t^\gamma}{\gamma}\left(1 - \dfrac{3}{2\,t} - \dfrac{1}{12\,t^2} + o\big(t^{-2}\big)\right) & \text{if } \gamma = 1 \\[2ex]
\dfrac{t^\gamma}{\gamma}\left(1 - \dfrac{\gamma}{2\,t} + o\big(t^{-1}\big)\right) & \text{if } \gamma > 1 .
\end{cases}
\end{aligned}$$

Then

$$U(t\,x) - U(t) \;=\; \begin{cases}
t^\gamma\left(\dfrac{x^\gamma - 1}{\gamma} - \dfrac{\gamma-1}{2\,t}\left(\dfrac{x^{\gamma-1}-1}{\gamma-1}\right) + o\big(t^{-1}\big)\right) & \text{if } \gamma \neq 1 \\[2ex]
t^\gamma\left(\dfrac{x^\gamma - 1}{\gamma} + \dfrac{\gamma-2}{12\,t^2}\left(\dfrac{x^{\gamma-2}-1}{\gamma-2}\right) + o\big(t^{-2}\big)\right) & \text{if } \gamma = 1 ,
\end{cases}$$

i.e., we may choose, in (2.3),

$$a_0(t) = t^\gamma \quad \text{and} \quad A_0(t) = \begin{cases} -\dfrac{\gamma-1}{2\,t} & \text{if } \gamma \neq 1 \\ -\dfrac{1}{12\,t^2} & \text{if } \gamma = 1 \end{cases}, \quad \text{with } \rho = \begin{cases} -1 & \text{if } \gamma \neq 1 \\ -2 & \text{if } \gamma = 1 . \end{cases}$$

Since

$$1 - A_0(t) = \begin{cases} \dfrac{2\,t+\gamma-1}{2\,t} & \text{if } \gamma \neq 1 \\ \dfrac{12\,t^2-1}{12\,t^2} & \text{if } \gamma = 1 , \end{cases}$$

we get, from (2.4),

$$a(t) = \frac{a_0(t)}{1-A_0(t)} = \begin{cases} \dfrac{2\,t^{\gamma+1}}{2\,t+\gamma-1} & \text{if } \gamma \neq 1 \\ \dfrac{12\,t^3}{12\,t^2-1} & \text{if } \gamma = 1 \end{cases}$$

and

$$A(t) = \rho\, A_0(t) = \begin{cases} \dfrac{\gamma-1}{2\,t} & \text{if } \gamma \neq 1 \\ \dfrac{1}{6\,t^2} & \text{if } \gamma = 1 . \end{cases}$$

Then

$$\frac{a(t)}{U(t)} = \begin{cases} -\gamma\, t^\gamma \left(1 + \left(\dfrac{1-\gamma}{2\,t} + t^\gamma\right)\big(1+o(1)\big)\right) & \text{if } \gamma < 0 \\ \dfrac{1}{\ln t}\left(1 + \dfrac{1}{2\,t} + o\big(t^{-1}\big)\right) & \text{if } \gamma = 0 \\ \gamma\Big(1 + t^{-\gamma} + o\big(t^{-\gamma}\big)\Big) & \text{if } 0 < \gamma < 1 \\ 1 + \dfrac{3}{2\,t} + o\big(t^{-1}\big) & \text{if } \gamma = 1 \\ \gamma\left(1 + \dfrac{1}{2\,t} + o\big(t^{-1}\big)\right) & \text{if } \gamma > 1 , \end{cases}$$

and consequently,

$$\overline{A}(t) = \frac{a(t)}{U(t)} - \gamma_+ = \begin{cases} -\gamma\, t^\gamma \big(1+o(1)\big) & \text{if } \gamma < 0 \\ \dfrac{1}{\ln t}\big(1+o(1)\big) & \text{if } \gamma = 0 \\ \gamma\, t^{-\gamma} + o\big(t^{-\gamma}\big) & \text{if } 0 < \gamma < 1 \\ \dfrac{3}{2\,t} + o\big(t^{-1}\big) & \text{if } \gamma = 1 \\ \dfrac{\gamma}{2\,t} + o\big(t^{-1}\big) & \text{if } \gamma > 1 . \end{cases}$$

Then

$$\frac{\overline{A}(t)}{A(t)} = \begin{cases} -\dfrac{2\,\gamma\,t^{\gamma+1}}{\gamma-1}\big(1+o(1)\big) & \text{if } \gamma<0 \\ -\dfrac{2\,t}{\ln t}\big(1+o(1)\big) & \text{if } \gamma=0 \\ \dfrac{2\,\gamma}{\gamma-1}\; t^{1-\gamma}\big(1+o(1)\big) & \text{if } 0<\gamma<1 \\ -9\,t\,\big(1+o(1)\big) & \text{if } \gamma=1 \\ \dfrac{\gamma}{\gamma-1}\big(1+o(1)\big) & \text{if } \gamma>1 \end{cases} \underset{t\to\infty}{\longrightarrow}$$

$$\underset{t\to\infty}{\longrightarrow} \begin{cases} 0 & \text{if } \gamma<-1 \\ -\infty & \text{if } -1<\gamma\le 1 \\ \dfrac{\gamma}{\gamma-1}=\dfrac{\gamma}{\gamma+\rho} & \text{if } \gamma>1 \end{cases} =:\; c\,.$$

Note that for $\gamma=\rho=-1$ we get a finite limit $\overline{A}(t)/A(t) \underset{t\to\infty}{\longrightarrow} -1$ and different from $\gamma/(\gamma+\rho)=1/2$.

Let us now compute for $0<\gamma<-\rho$,

$$U(t)-\frac{a(t)}{\gamma} = \begin{cases} \dfrac{t^\gamma}{\gamma}\Big(-t^{-\gamma}+o\big(t^{-\gamma}\big)\Big) & \text{if } 0<\gamma<1 \\ t\left(-\dfrac{3}{2\,t}+o\big(t^{-1}\big)\right) & \text{if } \gamma=1 \end{cases} \underset{t\to\infty}{\longrightarrow}$$

$$\underset{t\to\infty}{\longrightarrow} \begin{cases} -\dfrac{1}{\gamma} & \text{if } 0<\gamma<1 \\ -\dfrac{3}{2} & \text{if } \gamma=1 \end{cases} =:\; l\,,$$

in agreement with Theorem 2.1. Note however that $l\neq 0$ for all $0<\gamma<-\rho$ and $c=\pm\infty$ for all region $0<\gamma<-\rho$.

On another side, for heavy tails, i.e., for $\gamma>0$,

$$\frac{\ln U(tx)-\ln U(t)-\gamma\ln x}{\widetilde{A}(t)} \underset{t\to\infty}{\longrightarrow} \frac{x^{\widetilde{\rho}}-1}{\widetilde{\rho}}\,, \qquad \widetilde{\rho} = \begin{cases} -\gamma & \text{if } 0<\gamma\le 1 \\ \rho=-1 & \text{if } \gamma>1\,, \end{cases}$$

$$\widetilde{A}(t) = \begin{cases} \gamma\,t^{-\gamma} & \text{if } 0<\gamma<1 \\ \dfrac{3}{2\,t} & \text{if } \gamma=1 \\ \dfrac{\gamma}{2\,t} & \text{if } \gamma>1 \end{cases} = \overline{A}(t)\,\big(1+o(1)\big)\,,$$

now in agreement with the results in Corollary 3.1.

Example 4.4. The most common heavy-tailed models with $\widetilde{\rho} = -\gamma$ and $0 < \gamma < -\rho$ (then necessarily with $l \neq 0$), are such that

$$U(t) \;=\; C\,t^{\gamma}\Big(1 + A\,t^{-\gamma} + B\,t^{-2\gamma} + o\big(t^{-2\gamma}\big)\Big)\;, \qquad A, B \neq 0, \quad C > 0\;.$$

For these models,

$$U(tx) - U(t) \;=\; C\,\gamma\,t^{\gamma}\left(\frac{x^{\gamma}-1}{\gamma} - B\,t^{-2\gamma}\left(\frac{x^{-\gamma}-1}{-\gamma}\right) + o\big(t^{-2\gamma}\big)\right)\;,$$

and

$$\frac{\frac{U(tx)-U(t)}{C\,\gamma\,t^{\gamma}} - \frac{x^{\gamma}-1}{\gamma}}{-B\,t^{-2\gamma}} \;\underset{t\to\infty}{\longrightarrow}\; \frac{x^{-\gamma}-1}{-\gamma}\;,$$

i.e., $\rho + \gamma = -\gamma$, or equivalently, $\rho = -2\,\gamma$. Then, (2.2) holds, provided that we choose

$$a(t) \;=\; \frac{C\,\gamma\,t^{\gamma}}{1 + B\,t^{-2\gamma}}\;, \qquad A(t) \;=\; 2\,B\,\gamma\,t^{-2\gamma}$$

and

$$\frac{a(t)}{U(t)} \;=\; \gamma\Big(1 - A\,t^{-\gamma} - \big(2\,B - A^2\big)\;t^{-2\gamma} + o\big(t^{-2\gamma}\big)\Big)\;.$$

Consequently, with $\overline{A}(t)$, l and c provided in (2.5), (2.6) and (2.7), respectively,

$$\overline{A}(t) \;=\; -A\,\gamma\,t^{-\gamma}\,\big(1 + O(t^{-\gamma})\big)\;, \qquad \frac{\overline{A}(t)}{A(t)} \;=\; -\frac{A}{2\,B\,t^{-\gamma}}\,\big(1 + O(t^{-\gamma})\big) \underset{t\to\infty}{\longrightarrow} \pm\infty\;,$$

i.e., $c = \pm\infty$ and

$$U(t) - \frac{a(t)}{\gamma} \;=\; C\,t^{\gamma}\Big(A\,t^{-\gamma} + 2\,B\,t^{-2\gamma} + o\big(t^{-2\gamma}\big)\Big) \underset{t\to\infty}{\longrightarrow} AC \neq 0\;,$$

i.e., $l = AC \neq 0$, as mentioned at the very beginning of this example. Indeed, we could also have written

$$U(t) \;=\; l + C\,t^{\gamma}\Big(1 + B\,t^{-2\gamma} + o\big(t^{-2\gamma}\big)\Big), \qquad \text{as}\;\; t \to \infty\;.$$

5. THE SECOND ORDER CONDITION FOR A GENERAL TAIL, HEAVY TAIL AND A THIRD ORDER FRAMEWORK

Note that for heavy-tailed models, the second order condition (2.2) implies a third order behaviour of the function $\ln U(t)$, whenever we are in the region $0 < \gamma \leq -\rho$, and $l \neq 0$, a region where $A(t) = o\big(\overline{A}(t)\big)$. Also, since $\big|\overline{A}\big| \in RV_{-\gamma}$, $|A| \in RV_{\rho}$ and $\overline{A}^2 \in RV_{-2\gamma}$, then A dominates $\overline{A}^2$ if $\rho > -2\,\gamma$, but $\overline{A}^2$ dominates A if $\rho < -2\,\gamma$. From the Proof of Theorem 3.1, Case (i), the third order behaviour of $\ln U(t)$ may be written as

$$\lim_{t\to\infty} \frac{\frac{\ln U(tx) - \ln U(t) - \gamma\,\ln x}{\overline{A}(t)} - \frac{x^{-\gamma}-1}{-\gamma}}{\widetilde{B}(t)} \;=\; H_{\widetilde{\rho},\widetilde{\eta}}(x)\;,$$

where H is defined in (2.2),

$$\widetilde{B}(t) := \begin{cases} -\overline{A}(t) & \text{if} \quad 0 < \gamma < -\dfrac{\rho}{2} \\ \gamma \, \dfrac{A(t)}{\overline{A}(t)} & \text{if} \quad -\dfrac{\rho}{2} < \gamma < -\rho \end{cases}$$

and the second and third order parameters $\widetilde{\rho}$ and $\widetilde{\eta}$, respectively, are given by

$$\widetilde{\rho} = -\gamma \;, \qquad \widetilde{\eta} = \begin{cases} -\gamma & \text{if} \quad 0 < \gamma < -\dfrac{\rho}{2} \\ \gamma + \rho & \text{if} \quad -\dfrac{\rho}{2} < \gamma < -\rho \;. \end{cases}$$

Note that in the region $-\rho/2 < \gamma < -\rho$ we get $\widetilde{\rho} \neq \widetilde{\eta}$.

Remark 5.1. For the case $\gamma = -\rho/2$, excluded from this note, everything depends on the relative behaviour of A and $\overline{A}^2$, both regularly varying functions with the same index of regular variation ρ.

Note also that the situation $\widetilde{\eta} = \widetilde{\rho}$ is the one that most often happens in practice, for standard heavy-tailed models like Fréchet, Burr, the Generalized Pareto and Student's t d.f.'s. For these d.f.'s, (2.2) holds with $\rho = -2\,\gamma$. However, if we induce a shift $l \neq 0$ in the above mentioned models, this relation between γ and ρ no longer exists and we may cover all region $0 < \gamma < -\rho$.

Finally, we mention that for the extreme value model with d.f. G_γ, we get

$$\rho = -1 \;, \qquad \widetilde{\rho} = -\gamma \qquad \text{and} \qquad \widetilde{\eta} = \begin{cases} -\gamma & \text{if} \quad 0 \leq \gamma \leq 1/2 \\ \gamma - 1 & \text{if} \quad 1/2 < \gamma < 1 \;. \end{cases}$$

For more details on the way the third order framework may be used in Statistics of Extremes, see, for heavy tails, Gomes, de Haan and Peng (2002) and Fraga Alves, Gomes and de Haan (2003a), papers dealing with the estimation of the second order parameter ρ, and Gomes, Caeiro and Figueiredo (2004), a paper dealing with reduced bias extreme value index estimation. For details on the general third order framework, see Fraga Alves, de Haan and Lin (2003b, Appendix; 2006).

As a final remark, we would like to emphasise the importance of all these conditions in Statistics of Extremes. The first order conditions in (2.1), (2.8) and (2.9), together with additional light conditions on k, the number of top order statistics used in the estimation of a first order parameter, enable us to guarantee consistency of semi-parametric estimators of such a parameter. The primary first order parameter is the extreme value index γ, but we can refer other relevant parameters of extreme events, like high quantiles, return periods or probabilities of exceedances of high levels, among others. To obtain a Central Limit Theorem for these estimators, or consistency of any estimator of a second order parameter, like the shape second order parameters ρ or $\widetilde{\rho}$, discussed in this paper, it is convenient to assume a second order condition, like the ones in (2.2) and (2.10).

For the derivation of an asymptotic non-degenerate behaviour of estimators of second order parameters, we further need to assume a third order condition, ruling the rate of convergence in (2.2) or in (2.10). Such a type of condition is also quite useful for the study of any second-order reduced-bias estimators, particularly if we want to have information on the bias of such estimators. For details on this type of extreme value index estimators and the importance of third order conditions see, for instance, the most recent papers on the subject (Caeiro, Gomes and Pestana, 2005; Gomes, de Haan and Henriques Rodrigues, 2007b; Gomes, Martins and Neves, 2007c). In these papers, the adequate external estimation of second order parameters leads to reduced-bias estimators with the same asymptotic variance as the (biased) classical estimator for heavy tails, the Hill estimator (Hill, 1975). For overviews on second-order reduced-bias estimation see Reiss and Thomas (Chapter 6) and Gomes, Canto e Castro, Fraga Alves and Pestana (2007a).

ACKNOWLEDGMENTS

Research supported by FCT/POCI 2010 and POCI – ERAS Project MAT/58876/2004. We also acknowledge the valuable suggestions from the referees.

REFERENCES

[1] Caeiro, F.; Gomes, M.I. and Pestana, D. (2005). Direct reduction of bias of the classical Hill estimator, *Revstat*, **3**(2), 113–136.

[2] Draisma, G.; Haan, L. de; Peng, L. and Pereira, T.T. (1999). A bootstrap-based method to achieve optimality in estimating the extreme value index, *Extremes*, **2**(4), 367–404.

[3] Ferreira, A.; Haan, L. de and Peng, L. (2003). On optimizing the estimation of high quantiles of a probability distribution, *Statistics*, **37**(5), 401–434.

[4] Fraga Alves, M.I.; Gomes, M.I. and Haan, L. de (2003a). A new class of semi-parametric estimators of the second order parameter, *Portugaliae Mathematica*, **60**(2), 194–213.

[5] Fraga Alves, M.I.; Haan, L. de and Lin, T. (2003b). Estimation of the parameter controlling the speed of convergence in extreme value theory, *Math. Methods of Statist.*, **12**, 155–176.

[6] Fraga Alves, M.I.; Haan, L. de and Lin, T. (2006). Third order extended regular variation, *Publications de l'Institut Mathématique*, **80**(94), 109–120.

[7] Geluk, J. and Haan, L. de (1987). *Regular Variation, Extensions and Tauberian Theorems*, CWI Tract 40, Center for Mathematics and Computer Science, Amsterdam, The Netherlands.

[8] Gnedenko, B.V. (1943). Sur la distribution limite du terme maximum d'une série aléatoire, *Ann. Math.*, **44**, 423–453.

[9] Gomes, M.I.; Caeiro, F. and Figueiredo, F. (2004). Bias reduction of a tail index estimator trough an external estimation of the second order parameter, *Statistics*, **38**(6), 497–510.

[10] Gomes, M.I.; Canto e Castro, L.; Fraga Alves, M.I. and Pestana, D. (2007a). Statistics of extremes for IID data and breakthroughs in the estimation of the extreme value index: Laurens de Haan leading contributions, *Extremes*, in press.

[11] Gomes, M.I.; Haan, L. de and Henriques Rodrigues, L. (2007b). Tail index estimation for heavy-tailed models: accommodation of bias in the weighted log-excesses, *J. Royal Statistical Society*, **B**, in press.

[12] Gomes, M.I.; Haan, L. de and Peng, L. (2002). Semi-parametric estimation of the second order parameter — asymptotic and finite sample behaviour, *Extremes*, **5**(4), 387–414.

[13] Gomes, M.I.; Martins, M.J. and Neves, M. (2007c). Improving second order reduced bias extreme value index estimation, *Revstat*, **5**(2), 177–207.

[14] Gomes, M.I. and Neves, C. (2007). Asymptotic comparison of the mixed moment and classical extreme value index estimators, *Statistics and Probability Letters*, in press.

[15] Haan, L. de (1970). *On Regular Variation and its Application to Weak Convergence of Sample Extremes*, Mathematisch Centrum Amsterdam.

[16] Haan, L. de (1984). *Slow variation and characterization of domains of attraction.* In "Statistical Extremes and Applications" (Tiago de Oliveira, Ed.), D. Reidel, Dordrecht, 31–48.

[17] Hill, B.M. (1975). A simple general approach to inference about the tail of a distribution, *Ann. Statist.*, **3**(5), 1163–1174.

[18] Haan, L. de and Ferreira, A. (2006). *Extreme Value Theory: an Introduction*, Springer Series in Operations Research and Financial Engineering.

[19] Reiss, R.-D. and Thomas, M. (2007). *Statistical Analysis of Extreme Values, with Application to Insurance, Finance, Hydrology and Other Fields*, 3rd edition, Birkhäuser Verlag.

REVSTAT – Statistical Journal

Background

Statistical Institute of Portugal (INE), well aware of how vital a statistical culture is in understanding most phenomena in the present-day world, and of its responsibility in disseminating statistical knowledge, started the publication of the scientific statistical journal *Revista de Estatística*, in Portuguese, publishing three times a year papers containing original research results, and application studies, namely in the economic, social and demographic fields.

In 1998 it was decided to publish papers also in English. This step has been taken to achieve a larger diffusion, and to encourage foreign contributors to submit their work.

At the time, the Editorial Board was mainly composed by Portuguese university professors, being now composed by national and international university professors, and this has been the first step aimed at changing the character of *Revista de Estatística* from a national to an international scientific journal.

In 2001, the *Revista de Estatística* published three volumes special issue containing extended abstracts of the invited contributed papers presented at the 23rd European Meeting of Statisticians.

The name of the Journal has been changed to REVSTAT – STATISTICAL JOURNAL, published in English, with a prestigious international editorial board, hoping to become one more place where scientists may feel proud of publishing their research results.

— The editorial policy will focus on publishing research articles at the highest level in the domains of Probability and Statistics with emphasis on the originality and importance of the research.

— All research articles will be refereed by at least two persons, one from the Editorial Board and another, external.

— The only working language allowed will be English.

— Three volumes are scheduled for publication, one in March, one in June and the other in November.

— On average, four articles will be published per issue.

Aims and Scope

The aim of REVSTAT is to publish articles of high scientific content, in English, developing innovative statistical scientific methods and introducing original research, grounded in substantive problems.

REVSTAT covers all branches of Probability and Statistics. Surveys of important areas of research in the field are also welcome.

Abstract/indexed in

REVSTAT is expected to be abstracted/indexed at least in *Current Index to Statistics*, *Mathematical Reviews*, *Statistical Theory and Method Abstracts*, and *Zentralblatt für Mathematic*.

Instructions to Authors, special-issue editors and publishers

Papers may be submitted in two different ways:

— By sending a paper copy to the Executive Editor and one copy to one of the two Editors or Associate Editors whose opinion the author(s) would like to be taken into account, together with a postscript or a PDF file of the paper to the e-mail: revstat@fc.ul.pt.

— By sending a paper copy to the Executive Editor, together with a postscript or a PDF file of the paper to the e-mail: revstat@fc.ul.pt.

Submission of a paper means that it contains original work that has not been nor is about to be published elsewhere in any form.

Submitted manuscripts (text, tables and figures) should be typed only in black, on one side, in double spacing, with a left margin of at least 3 cm and not have more than 30 pages.

The first page should include the name, affiliation and address of the author(s) and a short abstract with the maximum of 100 words, followed by the key words up to the limit of 6, and the AMS 2000 subject classification.

Authors are obliged to write the final version of accepted papers using LaTeX, in the REVSTAT style.

This style (REVSTAT.sty), and examples file (REVSTAT.tex), which may be download to *PC Windows System* (Zip format), *Mackintosh*, *Linux* and *Solaris Systems* (StuffIt format), and *Mackintosh System* (BinHex Format), are available in the REVSTAT link of the National Statistical Institute's Website: http://www.ine.pt/revstat.html

Additional information for the authors may be obtained in the above link.

Accepted papers

Authors of accepted papers are requested to provide the LaTeX files and also a postscript (PS) or an acrobat (PDF) file of the paper to the Secretary of REVSTAT: liliana.martins@ine.pt.

Such e-mail message should include the author(s)'s name, mentioning that it has been accepted by REVSTAT.

The authors should also mention if encapsulated postscript figure files were included, and submit electronics figures separately in .tiff, .gif, .eps or .ps format. Figures must be a minimum of 300 dpi.

Also send always the final paper version to:

Maria José Carrilho
Executive Editor, REVSTAT – STATISTICAL JOURNAL
Instituto Nacional de Estatística
Av. António José de Almeida
1000-043 LISBOA
PORTUGAL

Copyright and Reprints

Upon acceptance of an article, the author(s) will be asked to transfer copyright of the article to the publisher, the INE, in order to ensure the widest possible dissemination of information, namely through the National Statistical Institute's Website (http://www.ine.pt).

After assigning the transfer copyright form, authors may use their own material in other publications provided that the REVSTAT is acknowledged as the original place of publication. The Executive Editor of the Journal must be notified in writing in advance.

Authors of articles published in the REVSTAT will be entitled to one free copy of the respective issue of the Journal and twenty-five reprints of the paper are provided free. Additional reprints may be ordered at expenses of the author(s), and prior to publication.