Estimating the Autocovariation Function with applications to Heavy-tailed ARMA modeling. [1]

By Colin M. Gallagher

Clemson University

We consider the *autocovariation function* which is an analogue of the autocorrelation function defined for stationary processes with finite absolute mean. The *sample autocovariation function* is shown to be a strongly consistent estimator of the autocovariation function, whenever the underlying process is stationary and ergodic. The limit distribution of the sample autocovariation function is given for both finite variance and infinite variance causal ARMA processes. Applications to time series modeling are given.

1. Introduction. Consider a stationary causal ARMA process with infinite order moving average representation,

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}; \qquad \sum_{j=0}^{\infty} j|\psi_j| < \infty, \tag{1.1}$$

where $\{Z_t\}$ is a mean zero independent identically distributed (iid) sequence with $E|Z_t| < \infty$ and such that $P(X_1 = 0) = 0$. If $E|Z_t|^2 = \infty$ we assume the

---

[1]Research supported in part by NAVY-ONR Grant N00014-93-10800

AMS 1991 *subject classifications.* Primary 62M10,60E07; secondary 90A20, 60G55

Key words and phrases. Heavy-tailed time series, autocovariation, ARMA process, linear dependence, model identification, diagnostics

sequence has regularly varying tail probabilities,

$$P(|Z_1| > x) = L(x)x^{-\alpha} \tag{1.2}$$

where $L(x)$ is slowly varying at ∞ and $\alpha \in (1,2)$, and

$$\frac{P(Z_1 > x)}{P(|Z_1| > x)} \to a \quad \text{and} \quad \frac{P(Z_1 < -x)}{P(|Z_1| > x)} \to b \tag{1.3}$$

as $x \to \infty$, $a \in [0,1]$ and $b = 1 - a$. The conditions (1.2) and (1.3) ensure the existence of a sequence of constants $\{a_n\}$ satisfying

$$nP(|Z_1| > a_n x) \to x^{-\alpha} \quad \text{for all} \quad x > 0. \tag{1.4}$$

It is easy to see that (1.3) and (1.4) imply that

$$nP(Z_1 > a_n x) \to a x^{-\alpha} \quad \text{and} \quad nP(Z_1 < -a_n x) \to b x^{-\alpha}. \tag{1.5}$$

For the process with $E|Z_1|^2 = \infty$ the autocorrelation function (ACF) is not defined. However, Davis and Resnick (1985,1986) examine the sample autocorrelation function $\hat{\rho}(k)$ for ARMA processes with infinite second moment and regularly varying tail probabilities. When the tail of the distribution of Z_1 is asymptotically equivalent to a Pareto with exponent $\alpha < 2$, the limiting distribution of $(n/\log n)^{1/\alpha}(\hat{\rho}(h) - \rho(h))$ is given by

$$\sum_{j=0}^{\infty} (|\rho(h+j) + \rho(h-j) - 2\rho(j)\rho(h)|^{\alpha})^{1/\alpha}\, U/V, \tag{1.6}$$

where V is $\alpha/2$-skewed stable and U has a symmetric α-stable distribution and is independent of V. When $EZ_1^2 < \infty$

$$\sum_{j=0}^{\infty} \left(|\rho(h+j) + \rho(h-j) - 2\rho(j)\rho(h)|^{2}\right)^{1/2} N, \tag{1.7}$$

and N has a normal distribution (stable with index $\alpha = 2$). This of course implies that $\hat{\rho}(h)$ converges to $\rho(h)$ in probability when the data is generated by an ARMA process.

As a time series modeling tool the sample ACF has some drawbacks when data is heavy tailed. These include

(i) The ACF is not defined for processes with infinite variance.

(ii) A discontinuity in the quantile of the limit distribution as $\alpha \to 2$.

(iii) A discontinuity in the normalization as $\alpha \to 2$.

(iv) For $\alpha < 2$ the limiting distribution provides a poor approximation for reasonable sample sizes, especially as α approaches 2.

(v) Cohen, Resnick and Samorodnitsky (1998) give examples of stationary ergodic process with $E|X| < \infty$ and $E|X^2| = \infty$ for which the sample ACF converges to a non-degenerate limit.

The discontinuity in the quantile of the limit distribution results from the divisor V which is large with positive probability when $\alpha < 2$, but becomes unity at $\alpha = 2$. The discontinuities of the quantile and normalization present a problem in application. For example, if $\alpha = 2$, but is estimated to be 1.98 the normalization and quantile used in a statistical application will be drastically wrong (see Figure 1 in Section 3). Simulation evidence in Runde (1998) and in Adler, Feldman, and Gallagher (1998) indicates that the large sample distribution of $\hat{\rho}(k)$ is not very useful for stable data. In fact in some cases a sample size on the order of one million is necessary in order to get an accurate large sample approximation (e.g. see Table 4 in Adler, Feldman and Gallagher).

In this paper we investigate the behavior of an analogue of the autocorrelation function which is well defined for stationary processes $\{X_t\}$ with $E|X_1| < \infty$.

Definition 1.1. For any zero mean stationary process with finite first absolute moment we define the *autocovariation function* (AcovF) by

$$\lambda(k) = \frac{EX_t S_{t-k}}{E|X_{t-k}|} \quad \textit{for} \quad k = 0, \pm 1, \pm 2, \ldots.$$

The *covariation* considered in Cambanis and Miller (1981) is a measure of dependence for symmetric *alpha*-stable (SαS) random vectors. For background on stable random processes and the covariation see Samorodnitsky and Taqqu (1994). When the process X_t has SαS finite dimensional distributions, $\lambda(k)$ is the normalized covariation of X_t on X_{t-k} and

$$EX_t S_{t-k} = \lambda(k) E|X_{t-k}|.$$

Given observations from a time series $X_1, \ldots, X_n$, the method of moments estimator of $\lambda(k)$ is

$$\hat{\lambda}(k) = \frac{\sum_{t=k+1}^{n} X_t S_{t-k}}{\sum_{t=1}^{n} |X_t|}, \tag{1.8}$$

for $k > 0$ with an obvious modification for $k < 0$. We refer to $\hat{\lambda}(k)$ as the *sample autocovariation function.*

We will see that $\hat{\lambda}(k)$ does not suffer from the drawbacks mentioned above whenever $E|X_t| < \infty$. However, if $\alpha \leq 1$ the AcovF is not defined and the sample AcovF can converge to a random limit even for iid data (See example 2.6).

The remainder of this paper is organized as follows. We state our main results in Section 2. We give some examples and statistical applications in Section 3. Finally we derive the weak limit of the sample AcovF in Section 4.

2. Main results. We describe the limiting behavior of $\hat{\lambda}(k)$ in the following two theorems. Our first result, which follows from the ergodic theorem, is that $\hat{\lambda}(k)$ is strongly consistent whenever the process $\{X_t\}$ is stationary and ergodic.

This class of processes includes the causal ARMA processes. In Theorem 2.2 we give the joint asymptotic distribution of the sample AcovF for causal ARMA processes.

THEOREM 2.1. *If $\{X_t\}$ is a stationary ergodic sequence with $E|X_1| < \infty$, then*

$$\hat{\lambda}(k) \xrightarrow{a.s.} \lambda(k).$$

THEOREM 2.2. *Let X_t satisfy (1.1). For any h let*

$$\boldsymbol{\lambda}_h = (\lambda(-h), \dots \lambda(h)) \quad \textit{and} \quad \hat{\boldsymbol{\lambda}}_h = (\hat{\lambda}(-h), \dots \hat{\lambda}(h)).$$

(i) If $E|Z_1|^2 < \infty$ then

$$n^{1/2}(\hat{\boldsymbol{\lambda}}_h - \boldsymbol{\lambda}_h) \Rightarrow (E|X_1|)^{-1}\mathbf{X},$$

where $\mathbf{X}$ is multivariate normal.

(ii) If Z_1 satisfies (1.2) and (1.3),

$$na_n^{-1}(\hat{\lambda}_h - \lambda_h) \Rightarrow (E|X_1|)^{-1}\mathbf{Y},$$

where $\mathbf{Y}$ has a multivariate stable distribution.

REMARK 2.3. *The limiting vector $\mathbf{Y}$ in Theorem 2.2 can be represented as*

$$\mathbf{W} \begin{pmatrix} Y_1 \\ \vdots \\ Y_{2^h} \end{pmatrix},$$

where $Y_1, \dots Y_{2^h}$ are iid skewed stable random variables and the matrix $\mathbf{W}$ is given below in Section 4.

Remark 2.4. *If the tails of the distribution of Z_1 are asymptotically equivalent to a Pareto, then the normalization becomes $n^{1-1/\alpha}$ and there is no discontinuity at $\alpha = 2$. Also, there will be no discontinuity in the quantile as $\alpha \to 2$.*

Using properties of stable distributions we obtain a corollary to Theorem 2.2.

Corollary 2.5. *Let $\{X_t\}$ be given by (1.1) with $Z_1 \sim S\alpha S$ for some $\alpha > 1$.*

$$n^{1-1/\alpha}(\hat{\lambda}(k) - \lambda(k)) \Rightarrow \sigma_k X$$

where X has a stable distribution,

$$\lambda(k) = \frac{\sum_{j=k}^{\infty} \psi_j \psi_{j-k}{}^{<\alpha-1>}}{\sum_{j=0}^{\infty} |\psi_j|^{\alpha}},$$

and

$$\sigma_k = (\Sigma_{j=0}^{\infty} |\psi_j|^{\alpha})^{-1/\alpha}.$$

In the case $\alpha < 1$, the sample ACF converges to a constant if $\{X_t\}$ is an ARMA process. As the next example shows, a similar result does not hold for the sample AcovF.

Example 2.6. *Let $\{X_t\}$ be an iid symmetric sequence satisfying (1.2) for some $\alpha < 1$. For any $k > 0$,*

$$\hat{\lambda}(k) \Rightarrow \frac{Y_1 + Y_2}{Y_1 - Y_2},$$

where Y_1 and Y_2 are independent skewed α-stable random variables.

3. Examples and Applications. In this section we consider some applications of Corollary 2.5.

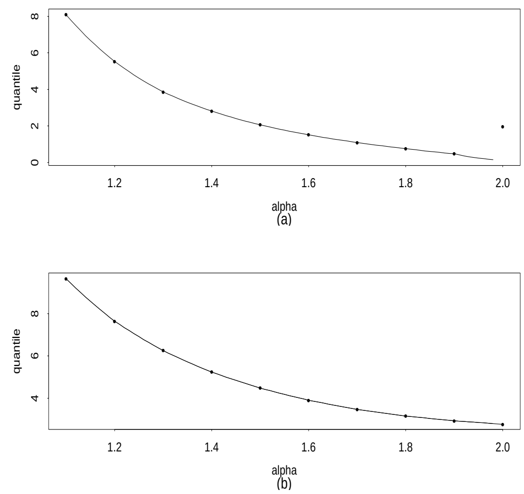

FIG. 1. *(a) .975 quantiles of U/V and N in (3.2) and (b) .975 quantiles of X in (3.3). Quantiles for alpha values other than* $1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9,$ *and* 2 *were approximated using interpolation.*

The SαS AR(1) model. Consider the AR(1) model:

$$X_t = \phi X_{t-1} + Z_t, \tag{3.1}$$

where $Z_t \sim S\alpha S$ with $\alpha > 1$. The parameter ϕ can be estimated with the sample ACF at lag 1 or the AcovF at lag 1 . The limiting distributions of the sample ACF and AcovF can be used to derive confidence intervals for ϕ based on their respective estimates.

For this model

$$\begin{aligned} (n/\log n)^{(1/\alpha)}(\hat\rho(1)-\phi) &\Rightarrow (1-\phi^\alpha)^{1/\alpha}U/V \quad \text{when } \alpha < 2, \\ n^{1/2}(\hat\rho(1)-\phi) &\Rightarrow (1-\phi^2)^{1/2}N \quad \text{when } \alpha = 2. \end{aligned} \tag{3.2}$$

For a fixed sample size n the normalizing constant has a discontinuity at $\alpha = 2$. Likewise, the limiting distribution has a discontinuity at $\alpha = 2$ (The limit as $\alpha \to 2$ in the characteristic function of U/V is not the characteristic function of the normal distribution). This can be seen from Figure 1 a) which graphs the .975 quantile of U/V and N in (3.2) versus α. While

$$n^{1-1/\alpha}(\hat{\lambda}(1) - \phi) \Rightarrow (1 - \phi^{\alpha})^{1/\alpha} X, \tag{3.3}$$

where $X \sim S\alpha S$ for $\alpha \in (1, 2]$. Note that there is no discontinuity in either the normalization or the limiting distribution.

Finite order moving average. We can use the sample AcovF or the sample ACF to make preliminary estimates of the order of a finite order SαS MA(q)

$$X_t = Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q},$$

where $Z_1 \sim$ SαS.

EXAMPLE 3.1. *Let $\{X_t\}$ be a $S\alpha S$ MA(q) process and define*

$$\mathbf{S} = (S_0, S_1, , \ldots, S_q)^t \quad \textit{and} \quad \theta = (\theta_0, \theta_1, \ldots, \theta_q)^t;$$

for $|k| > q$,

$$n^{1-1/\alpha}\hat{\lambda}(k) \Rightarrow \left(\frac{E|\mathbf{S}^t\boldsymbol{\theta}|^{\alpha}}{\sum_{j=0}^{q}|\theta_j|^{\alpha}}\right)^{1/\alpha} X \tag{3.4}$$

where $X \sim S\alpha S(\pi/(2\Gamma(1-1/\alpha)))$. Notice that the constant in front of X reduces to 1 for the iid process, since $q = 0$. If the process is Gaussian (3.4) becomes

$$n^{1/2}\hat{\lambda}(k) \Rightarrow \sqrt{\frac{\pi}{2}}\left(1 - \sum_{k=1}^{q} E(S_0 S_k)\rho(k)\right)^{1/2} N, \tag{3.5}$$

where N has a standard normal distribution and $\rho(k)$ is the ACF at lag k.

For the MA(q) process, (1.6) and (1.7) give for $k > q$

$$\begin{aligned} (n/\log n)^{(1/\alpha)}\hat{\rho}(k) &\Rightarrow \left(1 + 2\textstyle\sum_{j=1}^{q}|\rho(j)|^{\alpha}\right)^{1/\alpha} U/V \qquad \alpha < 2,\\ n^{1/2}\hat{\rho}(k) &\Rightarrow \left(1 + 2\textstyle\sum_{j=1}^{q}|\rho(j)|^{2}\right)^{1/2} N \qquad \alpha = 2. \end{aligned}$$

We can use the quantiles of U/V and N to attempt to identify q using the sample ACF or the quantiles of X to attempt to identify q using the sample AcovF. To compare these two procedures we ran two simulation studies. We considered three possible strategies that a practitioner might employ:

(i) The practitioner may be unaware that his data is coming from a process with heavy tails. In such a case he might plot $\hat{\rho}(k)$ at various lags and compare to the quantiles $1.96/\sqrt{n}$ of the normal distribution.

(ii) The practitioner might plot $\tilde{\rho}_\alpha = (n/\log n)^{(1/\alpha)}\hat{\rho}(k)$ at various lags and compare to the .025 and .975 quantiles of the distribution of U/V given in Adler, Feldman and Gallagher (1997).

(iii) The practitioner might plot $\hat{\lambda}(k)$ at various lags and compare to the .025 and .975 quantiles of $n^{1/\alpha-1}X$ (quantiles for stable distributions can be found in Samorodnitsky and Taqqu, 1994).

We simulated data under two different moving average models.

$$X_t = Z_t + (-0.8)Z_{t-1}, \tag{3.6}$$

and

$$X_t = Z_t + (-0.3)Z_{t-1} + (0.7)Z_{t-2}. \tag{3.7}$$

For each model and for 10 different values of $\alpha \in (1,2]$, we simulated $1,000$ time series each having sample size $n = 1,000$. For each time series we calculated both the sample ACF and sample AcovF for lags $1, 2, \ldots, 10$. We counted the number of times each strategy correctly identified the order q. For the MA(1) data we counted a correct identification if the appropriate statistic was outside its respective bound at lag 1 and within the bound for each lag $2, 3, \ldots, 10$. For

α	Statistic $\hat{\rho}(k)$	$\tilde{\rho_\alpha}(k)$	$\hat{\lambda}(k)$	α	Statistic $\hat{\rho}(k)$	$\tilde{\rho_\alpha}(k)$	$\hat{\lambda}(k)$
1.1	810	905	20	1.6	623	684	859
1.2	768	883	43	1.7	598	534	805
1.3	727	874	292	1.8	552	313	743
1.4	702	844	839	1.9	498	82	604
1.5	682	793	902	2.0	464	NA	515

TABLE 1

Number of simulated data sets out of 1000 that were correctly identified as coming from an MA(1) process using strategies (i)-(iii).

the MA(2) data we counted a correct identification if the statistic was outside its respective bound at lags 1 and 2, and within the bound for each lag $3, \ldots, 10$. The results of the simulation are shown in Tables 1 and 2.

We see that the procedure using the autocovariation function correctly identified more data sets than either of the other two procedures for all $\alpha \geq 1.5$. It is common practice to use the ACF to identify the order of a Gaussian MA(q) model (e.g. See section 9.2 of Brockwell and Davis, 1991). From Tables 1 and 2, we see that in the Gaussian case the sample AcovF correctly identified more data sets than the ACF.

We also note that if we were unaware that our data was coming from a stable distribution and we used the sample ACF and the normal bounds, we would do better for smaller α than we did when $\alpha = 2$. For example for the MA(1) model, when $\alpha = 1.2$ the procedure which uses the Gaussian bounds correctly identified

α	Statistic		
	$\hat{\rho}(k)$	$\tilde{\rho_\alpha}(k)$	$\hat{\lambda}(k)$
1.1	821	913	18
1.2	771	880	18
1.3	735	870	96
1.4	707	849	626
1.5	703	821	887

α	Statistic		
	$\hat{\rho}(k)$	$\tilde{\rho_\alpha}(k)$	$\hat{\lambda}(k)$
1.6	632	699	858
1.7	602	550	807
1.8	569	328	751
1.9	519	90	637
2.0	507	NA	543

TABLE 2

Number of simulated data sets out of 1000 that were correctly identified as coming from an MA(2) process using strategies (i)-(iii).

768 data sets, and for $\alpha = 2$ this same procedure (used in the traditional Box-Jenkins approach) correctly identified only 464 data sets.

4. Proofs. In this section we derive the weak limit given in Theorem 2.2. We will deal with the finite variance and the infinite variance process seperately. For the finite variance process $a_n = \sqrt{n}$.

A simple application of the ergodic theorem gives

$$n^{-1} \sum_{t=1}^{n} |X_t| \xrightarrow{a.s.} E|X_1|. \tag{4.8}$$

It is easy to see that as $n \to \infty$,

$$n a_n^{-1} \left[(\hat{\lambda}(k) - \lambda(k)) - \left(\sum_{t=1}^{n} |X_t| \right)^{-1} \sum_{t=1}^{n} S_t (X_{t+k} - \lambda(k) X_t) \right] \xrightarrow{P} 0, \tag{4.9}$$

where $\xrightarrow{P}$ denotes convergence in probability.

Weak limit under finite variance. For an ARMA process with finite variance standard mixing theory can be used to prove a central limit theorem.

For any integer h define the vector $\mathbf{Y}_t$ by

$$\mathbf{Y}_t = (S_t X_{t+h} - E S_t X_{t+h}, \ldots, S_t X_{t-h} - E S_t X_{t-h})^t.$$

The weak limit of the sample AcovF is determined by the vector

$$(\sqrt{n})^{-1} \sum_{t=1}^{n} \mathbf{Y}_t = (\sqrt{n})^{-1} (\sum_{t=1}^{n} S_t X_{t+h} - E S_t X_{t+h}, \ldots, \sum_{t=1}^{n} S_t X_{t-h} - E S_t X_{t-h})^t.$$

PROPOSITION 4.1.

$$(\sqrt{n})^{-1} \sum_{t=1}^{n} \mathbf{Y}_t \Rightarrow \mathbf{X},$$

where $\mathbf{X}$ *has a multivariate normal distribution with covariance matrix* $\mathbf{V}$ *given by (4.11) below.*

PROOF. It is enough to show that

$$(\sqrt{n})^{-1} \sum_{t=1}^{n} \mathbf{c}^t Y_t \Rightarrow \mathbf{c}^t \mathbf{X},$$

for any $\mathbf{c} \in \Re^{2h+1}$. For each $|k| \le h$ define

$$Y_{rt}^{(k)} = \sum_{i=0}^{r} \psi_j (Z_{t+k-j} S_t - E Z_{t+k-j} S_t),$$

and

$$Y_t^{(k)} = S_t X_{t+k} - E S_t X_{t+k}.$$

The $Y_t^{(k)}$ have mean zero and finite variance. We have

$$\sum_{r=1}^{\infty} \left(E(Y_0^{(k)} - Y_{0r}^{(k)})^2 \right)^{1/2} \le \sum_{r=1}^{\infty} (E|Z_1|^2)^{1/2} \left(\sum_{i=r+1}^{\infty} \psi_i^2 \right)^{1/2}$$
$$\le (E|Z_1|^2)^{1/2} \sum_{r=1}^{\infty} r|\psi_r|$$

For $\mathbf{Y}_{0r} = (Y_{0r}^{(-h)}, \ldots, Y_{tr}^{(h)})^t$,

$$\sum_{r=1}^{\infty} \left(E\mathbf{c}^t(\mathbf{Y}_0 - \mathbf{Y}_{0r})^2\right)^{1/2} \leq (\mathbf{c}^t\mathbf{c}EZ_1^2)^{1/2} \sum_{j=0}^{\infty} j|\psi_j| < \infty \tag{4.10}$$

Applying Theorem 19.3 from Billingsley (1999), condition (4.10) implies that

$$(\sqrt{n}\sigma)^{-1} \sum_{t=1}^{n} \mathbf{c}^t Y_t \Rightarrow N(0,1),$$

where

$$\sigma^2 = E(\mathbf{c}^t Y_0)^2 + 2\sum_{t=1}^{\infty} E(\mathbf{c}^t Y_0 \mathbf{c}^t Y_t).$$

We can represent σ^2 by

$$\mathbf{c}^t \mathbf{V} \mathbf{c},$$

where,

$$\mathbf{V} = E\left(\mathbf{Y}_0\mathbf{Y}_0^t + \left(\sum_{t=1}^{\infty} \mathbf{Y}_t\right)\mathbf{Y}_0^t + \mathbf{Y}_0\left(\sum_{t=1}^{\infty} \mathbf{Y}_t\right)^t\right). \quad \square \tag{4.11}$$

Proof of part (*i*) of Theorem 2.2. Notice that $E(X_t S_{t-k} - \lambda(k)|X_t|) = 0$ and let $\mathbf{A}$ be the $2^h \times 2^h + 1$ matrix

$$A = \begin{pmatrix} \mathbf{I}_h & \mathbf{c}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{c}_2 & \mathbf{I}_h \end{pmatrix},$$

where $\mathbf{I}_h$ is the $h \times h$ identitiy matrix, $\mathbf{0}$ is a matrix with all zero entries and $\mathbf{c}_1 = (-\lambda(h), \ldots, -\lambda(1))^t$ and $\mathbf{c}_2 = (-\lambda(-1), \ldots, -\lambda(-h))^t$. From (4.8), (4.9) and Proposition 4.1

$$na_n^{-1}(\hat{\boldsymbol{\lambda}}_h - \boldsymbol{\lambda}_h) \Rightarrow \mathbf{Y}$$

where $\mathbf{Y}$ has a zero mean multivariate normal distribution with covariance matrix

$$\mathbf{W} = (E|X_1|)^2 \mathbf{A}\mathbf{V}\mathbf{A}^t. \quad \square \tag{4.12}$$

Weak limit under infinite variance. For the process with infinite variance we first find the limit distribution of

$$\left(\sum_{t=1}^{n} Z_t S_{t-h} - EZ_t S_{t-h}, \ldots, \sum_{t=1}^{n} Z_t S_{t+m} - EZ_t S_{t+m}\right).$$

Let δ be independent of S_i for all i with $P(\delta = 1) = a$ and $P(\delta = -1) = b$, where a and b are given by (1.3). If $\psi_j = 0$ for $j > q$ define

$$\gamma^t = (\delta S_{-h}, \ldots, \delta S_{-1}, 1, \mathrm{sign}(\psi_1), \ldots, \mathrm{sign}(\psi_\mathrm{q}), \delta \mathrm{S}_{\mathrm{q}+1}, \ldots, \delta \mathrm{S}_\mathrm{m}),$$

otherwise let

$$\gamma^t = (\delta S_{-h}, \ldots, \delta S_{-1}, 1, \mathrm{sign}(\psi_1), \ldots, \mathrm{sign}(\psi_\mathrm{m})).$$

The random vector γ takes one of r values ($r = 2^h$ for the infinite order process). Let $\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \ldots \boldsymbol{\gamma}_r$ be an enumeration of these outcomes and $\pi_1, \ldots, \pi_r$ be their respective probabilities. The matrix $\mathbf{\Gamma}$ appearing in Propostition 4.2 below is given by

$$\mathbf{\Gamma} = (\pi_1^{1/\alpha}\boldsymbol{\gamma}_1, \pi_2^{1/\alpha}\boldsymbol{\gamma}_2, \ldots, \pi_r^{1/\alpha}\boldsymbol{\gamma}_r). \tag{4.13}$$

PROPOSITION 4.2.

$$a_n^{-1}\left(\sum_{t=1}^{n} Z_t S_{t-h} - EZ_t S_{t-h}, \ldots, \sum_{t=1}^{n} Z_t S_{t+m} - EZ_t S_{t+m}\right) \Rightarrow \mathbf{\Gamma}\mathbf{Y},$$

where $\mathbf{Y}$ *is given in remark (2.3).*

PROOF. We show that any for any $\mathbf{c} \in \Re^{h+m+1}$

$$a_n^{-1}\sum_{t=1}^{n}\left[Z_t \sum_{i=-h}^{m} c_i S_{t+i} - E\left(Z_t \sum_{i=-h}^{m} c_i S_{t+i}\right)\right] \Rightarrow \mathbf{c}^t\mathbf{\Gamma}\mathbf{Y}. \tag{4.14}$$

The sequence

$$Z_t^* = Z_t\left(\sum_{i=-h}^{m} c_i S_{t+i}\right) \tag{4.15}$$

is weakly dependent and can be shown to satisfy conditions D, D', and D'' of Davis (1983). The weak limit is determined by the tail behavior of the sequence which is given in terms of the vector $\boldsymbol{\gamma}$ in Lemma 4.4 which we state and prove at the end of this section. From Davis (1983) it follows that the left hand side of (4.14) converges to a stable random variable R with scale

$$\sigma = [E|\mathbf{c}^t\gamma|^\alpha]^{1/\alpha},$$

and skewness parameter

$$\beta = \frac{E(\mathbf{c}^t\boldsymbol{\gamma})^+ - E(\mathbf{c}^t\boldsymbol{\gamma})^-}{E|\mathbf{c}^t\boldsymbol{\gamma}|^\alpha}.$$

Using properties of stable random variables it is easy to see that

$$R \stackrel{d}{=} \sum_{i=1}^{r} \pi_i^{1/\alpha}\mathbf{c}^t\gamma_i Y_i,$$

where $Y_1, \ldots, Y_r$ are iid skewed stable random variables. □

For each $|k| \le h$, let

$$Y_{t,k}^{(m)} = Z_t \sum_{j=-|k|}^{m} (S_{t+j}(\psi_j - \lambda(k)\psi_{j-k}) - EZ_tS_{t+j}(\psi_j - \lambda(k)\psi_{j-k}) .$$

From Proposition 4.2 we have for every integer m

$$a_n^{-1}\sum_{t=1}^{n}\mathbf{Y}_t^{(m)} \Rightarrow \mathbf{W}_m\mathbf{Y},$$

where

$$\mathbf{Y}_t^{(m)} = \left(Y_{t,-h}^{(m)}, \ldots, Y_{t,h}^{(m)}\right).$$

To describe the matrix $\mathbf{W}_m$, let

$$\mathbf{s}^i = (s_{-h}^i, \ldots, s_{-1}^i)^t$$

be the i^{th} outcome of $(S_{-h}, \ldots, S_{-1})$ and let $l = \max(k, 0)$. Then the i^{th} entry of the k^{th} row of $\mathbf{W}_m$ is

$$\pi_i \left(\sum_{j=-l}^{-1} s_j^i (\psi_{j+k} - \lambda(k)\psi_j) + \sum_{j=0}^{m} \mathrm{sign}(\psi_j)(\psi_{\mathrm{j+k}} - \lambda(\mathrm{k})\psi_{\mathrm{j}}) \right). \tag{4.16}$$

Clearly

$$\mathbf{W}_m \mathbf{Y} \Rightarrow \mathbf{W}\mathbf{Y}, \tag{4.17}$$

where the entries of $\mathbf{W}$ come from letting $m \to \infty$ in (4.16).

PROPOSITION 4.3.

$$a_n^{-1} \sum_{t=1}^{n} S_t (X_{t+k} - \lambda(k) X_t) \Rightarrow \mathbf{W}\mathbf{Y}$$

PROOF. We show that

$$\limsup_{n\to\infty} P\left(a_n^{-1} | \sum_{t=1}^{n} S_t (X_{t+k} - \lambda(k) X_t) - Y_{t,k}^{(m)} | > \epsilon \right) \to 0, \tag{4.18}$$

as $m \to \infty$. Using Theorem 3.2 in Billingsley (1999) the result follows from (4.18) and (4.17).

To prove (4.18), let $c_j = (\psi_{j+k} - \lambda(k)\psi_j)$. From the definition of $\lambda(k)$,

$$\sum_{j=-|k|}^{\infty} c_j E Z_t S_{t+j} = 0,$$

so that

$$\sum_{j<m} c_j E S_1 Z_{1-j} = -\sum_{j=m}^{\infty} c_j E \left(S_1 Z_{1-j} I_{|Z_{1-j}|>a_n} \right) - \sum_{j=m}^{\infty} c_j E \left(S_1 Z_{1-j} I_{|Z_{1-j}|\le a_n} \right).$$

Using this fact we have,

$$\sum_{t=1}^{n} S_t (X_{t+k} - \lambda(k) X_t) - Y_{t,k}^{(m)} = U_n + V_n + R_n,$$

Where

$$U_n = \sum_{t=1}^{n} S_t \sum_{j=m}^{\infty} c_j Z_{t-j} I_{|Z_{t-j}|\le a_n} - n \sum_{j=m}^{\infty} c_j E \left(S_1 Z_{1-j} I_{|Z_{1-j}|\le a_n} \right),$$

$$V_n = \sum_{t=1}^{n} S_t \sum_{j=m}^{\infty} c_j Z_{t-j} I_{|Z_{t-j}|>a_n} - nE\left(S_1 \sum_{j=m}^{\infty} c_j Z_{t-j} I_{|Z_{t-j}|>a_n},\right)$$

and

$$R_n = \sum_{t=1}^{n} S_t \sum_{j<m} c_j Z_{t-j} - \sum_{t=1}^{n} Z_t \sum_{j<m} c_j S_{t+j}.$$

We have the following

$$E|a_n^{-1} V_n| \leq 2na_n^{-1} E\left(|Z_1| I_{|Z_1|>a_n}\right) \sum_{j=m}^{\infty} |c_j|,$$

and

$$E|a_n^{-1} R_n| \leq a_n^{-1} E|Z_1| \left[(1+|\lambda(k)|) \sum_{j=0}^{\infty} j|\psi_j| + k \sum_{j=0}^{\infty} |\psi_j|\right].$$

A slight modification of the argument in Davis and Hsing 1995 page 915 gives

$$Var(U_n) \leq 2na_n^{-2} E Z_1^2 I_{|Z_1|\leq a_n} \left(\sum_{j=m}^{\infty} |c_j|\right)^2.$$

Using Karamata's Theorem (e.g. see Feller 1971 page 283), the left hand side of (4.18) is bounded by

$$\left(\frac{2\alpha}{\epsilon^2(2-\alpha)}\right)^{1/2} \left(\sum_{j=m}^{\infty} |c_j|\right)^2 + \frac{2\alpha}{\epsilon(\alpha-1)} \sum_{j=m}^{\infty} |c_j|,$$

which converges to 0 as $m \to \infty$ as was required. □

Proof of Theorem 2.2 (*ii*) From Proposition 4.3, (4.8) and (4.9) it follows from Slutsky's theorem that

$$na_n^{-1}(\hat{\boldsymbol{\lambda}}_h - \boldsymbol{\lambda}_h) \Rightarrow (E|X_1|)^{-1} \mathbf{WY},$$

where $\mathbf{Y}$ is given in remark (2.3). □

LEMMA 4.4. *For Z_t^* given by (4.15),*

$$nP(|Z_t^*| > a_n x) \to \sigma^\alpha x^{-\alpha} \quad \textit{for all} \quad x > 0$$

and

$$\frac{P(Z_1^* > a_n x)}{P(|Z_1^*| > a_n x)} \to \frac{E(\mathbf{c}^t \boldsymbol{\gamma}^+)^\alpha}{E|\mathbf{c}^t \boldsymbol{\gamma}|^\alpha} \quad \textit{and} \quad \frac{P(Z_1^* < -a_n x)}{P(|Z_1^*| > a_n x)} \to \frac{E(\mathbf{c}^t \boldsymbol{\gamma}^-)^\alpha}{E|\mathbf{c}^t \boldsymbol{\gamma}|^\alpha}$$

PROOF. Let $s_1, \ldots, s_r$ denote the possible outcomes of $S = (S_{-h}, \ldots, S_{-1})^t$ and define the two sets

$$B_1 = \left\{ s_i : \sum_{j=-h}^{-1} c_j s_j + \sum_{j=0}^{m} c_j \mathrm{sign}(\psi_\mathrm{j}) > 0 \right\}$$

and

$$B_2 = \left\{ s_i : \sum_{j=-h}^{-1} c_j s_j - \sum_{j=0}^{m} c_j \mathrm{sign}(\psi_\mathrm{j}) < 0 \right\}.$$

Let Z_0^* be given by (4.15), let $\delta_0 = \mathrm{sign}(\mathrm{Z}_0)$ and define the event:

$$A = \{S_j = \delta_0 \mathrm{sign}(\psi_\mathrm{j}) \quad \forall \mathrm{j} \quad \text{such that} \quad \psi_\mathrm{j} \neq 0\}.$$

We have

$$(4.19) \quad nP(Z_0^* > a_n x) = nP\left\{(Z_0^* > a_n x) \cap A\right\} + nP\left\{(Z_0^* > a_n x) \cap A^c\right\}.$$

We show that the second term in the right hand side of (4.19) converges to zero as $n \to \infty$.

For $c = \sum |c_i|$,

$$\begin{aligned} nP[Z_0^* > a_n x \cap A^c] &\le nP[|Z_0| > a_n x/c \cap A^c] \\ &\le n \sum_{j=0}^{m+h} P[Z_0 > a_n x^*, S_{1+j} = -\mathrm{sign}(\psi_\mathrm{j})] \\ &\quad + n \sum_{j=0}^{m+h} P[Z_0 < -a_n x^*, S_{1+j} = \mathrm{sign}(\psi_\mathrm{j})], \end{aligned}$$

where $x^* = x/c$. Each term in the above converges to zero as $n \to \infty$. We prove this for a term in the first sum; the other terms are handled similarly. Using elementary probability arguments, Markov's inequality and (1.4) we have for $\Gamma_i = \psi_i \frac{Z_{j-i}}{\psi_j}$,

$$\begin{aligned} nP[Z_0 > a_n x^*, S_j = -\mathrm{sign}(\psi_j)] &= nP\left[Z_0 > a_n x^*, Z_0 < -\sum_{i \neq j} \Gamma_i\right] \\ &\leq nP(Z_0 > a_n x^*)(a_n x^*)^{-1} \frac{E|Z_0|}{|\psi_j|} \sum_{i=0}^{\infty} |\psi_i| \\ &\to 0 \end{aligned}$$

as $n \to \infty$, since $(a_n)^{-1} \to 0$. Using the fact that Z_0 is independent of S_i for $i < 0$ and (1.5) we have:

$$\begin{aligned} nP\{(Z_0^* > a_n x) \cap A\} &= \sum_{B_1} \pi_i nP(Z_0 > a_n x |\mathbf{c}^t \boldsymbol{\gamma}_i|^{-1}) \\ &\quad + \sum_{B_2} \pi_i nP(Z_0 < -a_n x |\mathbf{c}^t \boldsymbol{\gamma}_i|^{-1}) \\ &\to \sum_{B_1} a\pi_i x^{-\alpha} |\mathbf{c}^t \boldsymbol{\gamma}_i|^{\alpha} + \sum_{B_2} b\pi_i x^{-\alpha} |\mathbf{c}^t \boldsymbol{\gamma}_i|^{\alpha} \\ &= x^{-\alpha} E((\mathbf{c}^t \boldsymbol{\gamma})^+)^{\alpha} \end{aligned}$$

A similar argument shows

$$nP(Z_0^* < -a_n x) \to x^{-\alpha} E((\mathbf{c}^t \boldsymbol{\gamma})^-)^{\alpha}.$$

Acknowledgments. I would like to thank Raya Feldman for reading early versions of this paper and an anonymous referee for his helpful comments.

REFERENCES

ADLER, R., FELDMAN. R., AND GALLAGHER, C. (1998). Analysing stable time series. In *A Practical Guide to Heavy Tails.* Birkhäuser, Boston.

Billingsley, P. (1999). *Convergence of Probability Measures.* Wiley, New York.

Brockwell, P. J. and Davis, R. A. (1991). *Time Series: Theory and Methods.* Springer, New York.

Cambanis, S. and Miller, G. (1981). Linear problems in pth order and stable processes. *SIAM J. Appl. Math.* **41**: 43-69.

J. Cohen, S. Resnick and G. Samorodnitsky (1997). Sample correlations of infinite variance time series models: an empirical and theoretical study. Available at www.orie.cornell.edu/trlist/trlist.html.

Davis, R. (1983). Stable limits for partial sums of dependent random variables. *Ann. Probab.* **11**: 262-269.

Davis, R. and Hsing, T. (1983). Point process and partial sum convergence for weakly dependent random variables with infinite variance. *Ann. Probab.* **23**: 879-917.

Davis, R. and Resnick, S. (1985). Limit theory for moving averages of random variables with regularly varying tail probabilities. *Ann. Probab.* **13**: 179-195.

Davis, R. and Resnick, S. (1986). Limit theory for the sample covariance and correlation functions of moving averages. *Ann. Statist.* **14**: 533–558.

Runde, R. (1998) The asymptotic null distribution of the Box-Pierce Q-statistic for random variables with infinite variance. *J. of Econometrics* **78**: 205-216.

Samorodnitsky, G. and Taqqu, M. (1994). *Stable Non-Gaussian Random Processes.* Chapman-Hall, New York.

DEPARTMENT OF MATHEMATICAL SCIENCES
CLEMSON UNIVERSITY
CLEMSON, SC 29634
CGALLAG@CLEMSON.EDU