Stable non-Gaussian time series

Ryukyu Math. J., 15(2002), 1-18

Stable non-Gaussian time series

CHUNHANG CHEN

1. Introduction

Classical time series analysis is mainly concerned with the statistical analysis of stationary linear processes

$$X_t = \sum_{j=-\infty}^{\infty} \psi_j Z_{t-j}, \ t \in \mathbb{Z},$$

with iid innovations $\{Z_t\}$ whose means are 0 and variances *finite*. The typical models include ARMA and fractionally integrated ARMA (FARIMA) models. In practical situations, however, many data exhibiting large fluctuations have been observed, which suggests that the marginal distributions are non-Gaussian and have heavy tails, so heavy that the distributions have infinite variance. In the last decade, there have been a rapid development in the statistical analysis of time series with infinite variance. In this report we provide a survey of these studies and give some new results for the least squares estimator (LSE) and the least absolute deviations (LAD) estimator of a time series regression model where the regression error term has infinite variance.

2. Stable distributions and domain of attraction

As a non-Gaussian distribution with infinite variance, the stable distributions have often been considered. A random variable X is said to have an α-stable distribution if there are parameters $0 < \alpha \leq 2$, $\sigma > 0$, $-1 \leq \beta \leq 1$ and $-\infty < \mu < \infty$ such that its characteristic function has the following form

$$E \exp i\theta X = \begin{cases} \exp\{-\sigma^\alpha|\theta|^\alpha(1 - i\beta(\operatorname{sign}\theta)\tan\frac{\pi\alpha}{2}) + i\mu\theta\} & \text{if } \alpha \neq 1, \\ \exp\{-\sigma|\theta|(1 + i\beta\frac{2}{\pi}(\operatorname{sign}\theta)\ln|\theta|) + i\mu\theta\} & \text{if } \alpha = 1, \end{cases}$$

where $\theta \in \mathbb{R}$ and

$$\operatorname{sign}\theta = \begin{cases} 1 & \text{if } \theta > 0, \\ 0 & \text{if } \theta = 0, \\ -1 & \text{if } \theta < 0. \end{cases}$$

Received November 30, 2002.

The parameter α is called the index of stability of the stable distribution, σ the scale parameter, β the skewness parameter and μ the shift parameter. We write this distribution as $S_\alpha(\sigma,\ \beta,\ \mu)$. When $\beta = 0$ and $\mu = 0$, the distribution $S_\alpha(\sigma, 0, 0)$ is called symmetric α-stable (SαS). An SαS rv X has a characteristic function of the simple form $E \exp i\theta X = e^{-\sigma^\alpha |\theta|^\alpha}$. When $\alpha = 2$, $S_2(\sigma, 0,\ \mu)$ is the normal distribution $N(\mu, 2\sigma^2)$; when $\alpha = 1$, $S_1(\sigma, 0,\ \mu)$ is a Cauchy distribution; when $\alpha = 1/2$, $S_{1/2}(\sigma, 1, 0)$ is the Lévy distribution. Except for these special cases, the probability densities of stable distributions are not known in explicit form. This makes it impossible to estimate parameters via MLE for a time series model in which the marginal distribution is stable.

If $X \sim S_\alpha(\sigma,\ \beta,\ \mu)$, $0 < \alpha < 2$, then X has the tail probabilities

$$\begin{cases} \lim_{\lambda\to\infty} \lambda^\alpha P\{X > \lambda\} = C_\alpha \frac{1+\beta}{2}\sigma^\alpha, \\ \lim_{\lambda\to\infty} \lambda^\alpha P\{X < -\lambda\} = C_\alpha \frac{1-\beta}{2}\sigma^\alpha, \end{cases}$$

where

$$C_\alpha = \left(\int_0^\infty x^{-\alpha} \sin x dx\right)^{-1} = \begin{cases} \frac{1-\alpha}{\Gamma(2-\alpha)\cos(\pi\alpha/2)} & \text{if } \alpha \neq 1, \\ 2/\pi & \text{if } \alpha = 1. \end{cases}$$

Note that $E|X|^r = \displaystyle\int_0^\infty P\{|X|^r > \lambda\}d\lambda$. From the tail behavior of stable distributions, we see that if $X \sim S_\alpha(\sigma,\ \beta,\ \mu)$, $0 < \alpha < 2$, then $E|X|^\alpha = +\infty$ and $E|X|^r < +\infty$ for $0 \le r < \alpha$. The fact that α-stable distributions with $\alpha < 2$ have infinite second moment means that many of the techniques valid for the Gaussian or finite-second-moment case do not apply. When $\alpha \le 1$, one also has $E|X| = +\infty$, precluding the use of expectations. These are the main source of difficulties when dealing the infinite-second-moment-case.

A sequence of iid rv's $\{Y_n\}$ is said to belong to the domain of attraction of an α-stable distribution G_α if there exist constants $\{a_n\}$ and $\{b_n\}$, $b_n > 0$, such that

$$\frac{Y_1 + \cdots + Y_n}{b_n} + a_n \Rightarrow G_\alpha$$

as $n \to \infty$, where $\Rightarrow$ denotes convergence in distribution. In general, $b_n = n^{1/\alpha} l(n)$, where $l(\cdot)$ is a slowly varying function at infinity, that is, $\lim_{x\to\infty} l(ux)/l(x) = 1$ for all $u > 0$. In particular, if we can choose $b_n = cn^{1/\alpha}$, where c is a constant, such that $(Y_1 + \cdots + Y_n)/b_n + a_n \Rightarrow G_\alpha$, then $\{Y_n\}$ is said to belong to the domain of normal attraction of an α-stable distribution G_α (write $\{Y_n\} \in \mathrm{DNA}(\alpha)$). It is well known that $\{Y_n\} \in \mathrm{DNA}(2)$ if and only if $EY_1^2 < \infty$; $\{Y_n\} \in \mathrm{DNA}(\alpha)$, $0 < \alpha < 2$, if and only if $P\{Y_1 < -\lambda\} \sim c_1\lambda^{-\alpha}$ and $P\{Y_1 > \lambda\} \sim c_2\lambda^{-\alpha}$ as $\lambda \to \infty$ for non-negative constants c_1, c_2 such that $c_1 + c_2 > 0$. For more details, see Feller (1971); Ibragimov and Linnik (1971).

3. How to measure the bivariate dependence for a time series with infinite variance?

The covariance function is an extremely powerful tool in time domain analysis of time series which is Gaussian or has finite variance, but it is not defined in the α-stable case when $\alpha < 2$. How do we measure the bivariate dependence for a time series with infinite variance? To do this, we need to consider the bivariate marginal distributions.

We introduce some definitions and notations. A d-dimensional random vector $\mathbf{X} = (X_1, \ldots, X_d)'$ in $\mathbb{R}^d$ is said to be symmetric α-stable (the rv's $X_1, \ldots, X_n$ is said to be jointly $S\alpha S$) if each linear combination $\sum_{k=1}^{d} b_k X_k$ is $S\alpha S$. A necessary and sufficient condition for $\mathbf{X}$ to be $S\alpha S$ is that there exists unique symmetric finite measure Γ on the unit sphere S_d in $\mathbb{R}^d$ (i.e., $\Gamma(-A) = \Gamma(A)$ for any Borel set A of S_d) such that

$$E\exp\{i(\boldsymbol{\theta}, \mathbf{X})\} = \exp\left\{-\int_{S_d} |(\boldsymbol{\theta}, \mathbf{X})|^{\alpha}\Gamma(ds)\right\}.$$

Here $(\cdot,\cdot)$ is the usual inner product on $\mathbb{R}^d$. Γ is called the spectral measure of $\mathbf{X}$. Let a and p be real numbers. Define the signed power $a^{<p>}$ as

$$a^{<p>} := |a|^p \text{sign } a = \begin{cases} a^p & \text{if } a \geq 0, \\ -|a|^p & \text{if } a < 0. \end{cases}$$

Let X_1 and X_2 be jointly $S\alpha S$ with spectral measure Γ. When $\alpha > 1$, define the covariation of X_1 and X_2 as

$$[X_1, X_2]_\alpha := \int_{S_2} s_1 s_2^{<\alpha-1>}\Gamma(ds),$$

where $\mathbf{s} = (s_1, s_2)' \in S_2$. When $\alpha = 2$, it holds that

$$[X_1, X_2]_\alpha = \frac{1}{2}\text{Cov}(X_1, X_2).$$

The covariation is not defined for $0 < \alpha \leq 1$. One can use another measure of bivaiate dependence, called the codifference. Let X_1 and X_2 be jointly $S\alpha S$, $0 < \alpha \leq 2$. Let σ_{X_1} denote the scale parameter of X_1. The codifference of X_1 and X_2 equals

$$\tau_{X_1,X_2} := \sigma_{X_1}^{\alpha} + \sigma_{X_2}^{\alpha} - \sigma_{X_1-X_2}^{\alpha}.$$

The codifference is defined for each $\alpha \in (0, 2]$. When $\alpha = 2$, we have

$$\tau_{X_1,X_2} = \text{Cov}(X_1, X_2).$$

The covariation and codifference share some of the properties of the covariance, but not all — many strange facts occur when $\alpha < 2$.

Whether or not the covariation and codifference play the same role for statistical analysis of time series with infinite variance as the covariance does in the finite-variance case has not been known. Furthermore, how to estimate the covariations and codifferences from a finite sample has not been known at all. Many problems deserve future studies.

4. Asymptotic behavior of the sample autocorrelation function

Consider a heavy-tailed stationary linear process $\{X_t\}$:

$$X_t = \sum_{j=-\infty}^{\infty} \psi_j Z_{t-j}. \tag{4.1}$$

We assume that $\{Z_t\}$ is a sequence of iid $S_\alpha(\sigma, 0, 0)$ rv's. That is, $E \exp i\theta Z_1 = e^{-|\sigma|^\alpha |\theta|^\alpha}$. The necessary and sufficient for the series (4.1) converges a.s. is that

$$\sum_{j=-\infty}^{\infty} |\psi_j|^\alpha < \infty, \tag{4.2}$$

and we assume that this condition is satisfied. As in the classical case, the sample autocorrelation function is defined by

$$\tilde{\rho}(h) := \frac{\sum_{t=1}^{n-|h|} X_t X_{t+|h|}}{\sum_{t=1}^{n} X_t^2}, \quad 0 \leq h < n.$$

Let

$$\rho(h) := \frac{\sum_{j=-\infty}^{\infty} \psi_j \psi_{j+|h|}}{\sum_{j=-\infty}^{\infty} \psi_j^2}, \quad h \in \mathbb{Z}.$$

In the classical case, $\{\rho(h)\}$ is the autocorrelation function; but in the stable case this can not be interpreted as autocorrelation function. However, we have for each $m \geq 1$,

$$\left(\frac{n}{\ln n}\right)^{1/\alpha} (\tilde{\rho}(h) - \rho(h))_{h=1,\dots,m} \Rightarrow (Y_h)_{h=1,\dots,m}, \tag{4.3}$$

where

$$Y_h := \sum_{j=1}^{\infty} [\rho(h+j) + \rho(h-j) - 2\rho(j)\rho(h)] \frac{G_j}{G_0}, \quad h = 1, \dots, m, \tag{4.4}$$

and $\{G_j\}_{j \geq 0}$ are independent stable rv's, G_0 is positive $\alpha/2$-stable with chf

$$E \exp i\theta G_0 = \exp\left\{-\Gamma(1-\alpha/2)\cos(\pi\alpha/4)|\theta|^{\alpha/2}(1 - i \text{ sign } (\theta) \tan(\pi\alpha/4))\right\}$$

and $\{G_j\}_{j\geq 1}$ are iid $S\alpha S$ rv'S with chf

$$E\exp i\theta G_1 = \begin{cases} \exp\{-\Gamma(1-\alpha)\cos(\pi\alpha/2)|\theta|^{\alpha}\} & \text{if } \alpha \neq 1, \\ \exp\{-\pi|\theta|/2\} & \text{if } \alpha \neq 1. \end{cases}$$

Notice that, $(\tilde{\rho}(h)-\rho(h))_{h=1,\dots,m}$ are asymptotically independently (normally) distributed in the classical case, but this no longer holds in the stable case. For details, see Davis and Resnick (1985, 1986).

The above results show that, although the autocorrelation function does not exist in the stable case, the sample autocorrelation function plays a similar role in the analysis of heavy-tailed time series as in the classical case, for example, in model-specification.

5. Asymptotic behavior of periodogram and integrated periodogram

Let $\{X_t\}$ be the heavy-tailed stationary linear process given in (4.1). Define the periodogram as

$$I_{n,X}(\lambda) := n^{-2/\alpha}\left|\sum_{t=1}^{n} X_t e^{-i\lambda t}\right|^2, \quad \lambda \in [-\pi,\pi].$$

Define the power transfer function of $\{X_t\}$ as

$$|\psi(\lambda)|^2 := \left|\sum_{j=-\infty}^{\infty} \psi_j e^{-i\lambda j}\right|^2, \quad \lambda \in [-\pi,\pi].$$

Assume that $|\psi(\lambda)|^2 > 0$ for any $\lambda \in [-\pi,\pi]$. Then for any frequencies $0 < \lambda_1 < \cdots < \lambda_m < \pi$, we have

$$(I_{n,X}(\lambda_h))_{h=1,\dots,m} \Rightarrow \left(|\psi(\lambda_h)|^2(\alpha^2(\lambda_h)+\beta^2(\lambda_h))\right)_{h=1,\dots,m},$$

where $(\alpha(\lambda_h),\ \beta(\lambda_h))_{h=1,\dots,m}$ is an $S\alpha S$ random vector in $\mathbb{R}^{2m}$, in which any two components are dependent. Moreover, if $\sum_{j=-\infty}^{\infty}|\psi_j| < \infty$, then

$$I_{n,X}(0) = (n^{-1/\alpha}\sum_{t=1}^{n} X_t)^2 \Rightarrow |\psi(0)|^2 S^2,$$

where S is an $S\alpha S$ rv.

Notice that, $(I_{n,X}(\lambda_h))_{h=1,\dots,m}$ are asymptotically independently (exponentially) distributed in the classical case, but this no longer holds in the stable case. For details, see Klüppelberg and Mikosch (1993).

Next consider the integrated periodogram

$$\int_{-\pi}^{x} I_{n,X}(\lambda) f(\lambda) d\lambda, \quad x \in [-\pi, \pi],$$

where $f(\cdot)$ is a smooth weight function. Define the self-normalised periodogram by

$$\tilde{I}_{n,X}(\lambda) := I_{n,X}(\lambda) / \Big(n^{-2/\alpha} \sum_{t=1}^{n} X_t^2 \Big), \quad \lambda \in [-\pi, \pi].$$

Let $f(\cdot)$ be a 2π-periodic continuous function such that the Fourier coefficients of $|f(\cdot)||\psi(\cdot)|^2$ are absolutely summable. Then we have

$$\int_{-\pi}^{x} I_{n,X}(\lambda) f(\lambda) d\lambda \Rightarrow Y_0 \int_{-\pi}^{x} |\psi(\lambda)|^2 f(\lambda) d\lambda \quad \text{in } \mathrm{C}[-\pi, \pi]$$

as $n \to \infty$, where $\mathrm{C}[-\pi, \pi]$ is the space of continuous functions on $[-\pi, \pi]$ equiped with the uniform topology, and Y_0 is an $\alpha/2$-stable positive rv which has the Laplace transform $E \exp(-\theta Y_0) = \exp\{-|\sigma|^\alpha K_\alpha \theta^{\alpha/2}\}$, where $K_\alpha = E|N|^{\alpha/2}$ and $N \sim N(0, 2)$. Now define

$$T_n := \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{I_{n,X}(\lambda)}{|\psi(\lambda)|^2} d\lambda.$$

If we restrict $\alpha \in (1, 2)$ and assume that $g(\lambda) := |\psi(\lambda)|^2 f(\lambda)$ is continuously differentiable, we have

$$\left(\frac{n}{C_\alpha \ln n} \right)^{1/\alpha} \int_{-\pi}^{x} (I_{n,X}(\lambda) - |\psi(\lambda)|^2 T_n) f(\lambda) d\lambda \Rightarrow \sigma^2 \left[g(x) S(x) - \int_{-\pi}^{x} g'(\lambda) S(\lambda) d\lambda \right]$$

in $\mathrm{C}[-\pi, \pi]$ as $n \to \infty$, where C_α is the same as given before,

$$S(x) := 2 \sum_{t=1}^{\infty} \frac{\sin(tx)}{t} Y_t,$$

where $\{Y_t\}_{t \geq 1}$ is a sequence of iid $S_\alpha(2^{-1/\alpha}, 0, 0)$ rv's with chf $E \exp i\theta Y_1 = e^{-|\theta|^\alpha/2}$ and is independent of Y_0. Note that, for $\alpha = 2$, $S(x)$ is a Brownian bridge on $[-\pi, \pi]$, see Hida (1980).

The above results are useful for deriving the asymptotic distributions of statistics in goodness-of-fit test for stable time series. For example,

Grenander-Rosenblatt test:

$$\left(\frac{n}{C_\alpha \ln n} \right)^{1/\alpha} \sup_{-\pi \leq x \leq \pi} \left| \int_{-\pi}^{x} (I_{n,X}(\lambda) - T_n) d\lambda \right| \Rightarrow \sigma^2 \sup_{-\pi \leq x \leq \pi} |S(x)|.$$

Cramér-von Mises test:

$$\left(\frac{n}{C_\alpha \ln n} \right)^{2/\alpha} \int_{-\pi}^{\pi} \left(\int_{-\pi}^{x} \left(\frac{I_{n,X}(\lambda)}{|\psi(\lambda)|^2} - T_n \right) d\lambda \right)^2 dx \Rightarrow \sigma^2 \int_{-\pi}^{\pi} S^2(x) dx = 4\pi\sigma^2 \sum_{t=1}^{\infty} \frac{Y_t^2}{t^2}.$$

For details, see Klüppelberg and Mikosch (1996).

6. Parameter estimation for heavy-tailed stationary ARMA and FARIMA models

Consider the stationary heavy-tailed ARMA(p,q) process:

$$X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q}, \qquad t \in \mathbb{Z}. \qquad (6.1)$$

Put

$$\beta = (\phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q)'$$

and let β_0 be the true, but unkown, parameter vector. Assume that $\phi(z,\beta) = 1 - \phi_1 z - \cdots - \phi_p z^p \neq 0$ and $\theta(z,\beta) = 1 + \theta_1 z + \cdots + \theta_q z^q \neq 0$ for $|z| \le 1$, and that, $\phi(z,\beta)$ and $\theta(z,\beta)$ have no common zeros, and let C be the parameter space consisting of β such that all these conditions are satisfied. Put

$$\psi(z,\beta) := \frac{\theta(z,\beta)}{\phi(z,\beta)} =: \sum_{j=0}^{\infty} \psi_j(\beta) z^j, \quad |z| \le 1.$$

Assume that $\{Z_t\}$ is a sequence of iid $S_\alpha(\sigma,0,0)$ rv's. Then (6.1) has a unique solution $X_t = \sum_{j=0}^{\infty} \psi_j(\beta) Z_{t-j}$ which converges a.s.

Now consider estimation of β_0. Since the explicit form of the probability density of $\{Z_t\}$ is in general unknown, it is difficult to construct the MLE for β_0. Here we consider the Whittle estimator:

$$\tilde{\beta}_n := \arg\min_{\beta \in C} \tilde{\sigma}_n^2(\beta),$$

where

$$\tilde{\sigma}_n^2 := \int_{-\pi}^{\pi} \frac{\tilde{I}_{n,X}(\lambda)}{|\psi(e^{i\lambda},\beta)|^2} d\lambda,$$

where $\tilde{I}_{n,X}(\lambda)$ is the selfnormalised periodogram. For practical purposes, the following version of the Whittle estimator is more appropriate: define the discretised version of $\tilde{\sigma}_n^2$ as

$$\hat{\sigma}_n^2 := \frac{2\pi}{n} \sum_{\lambda_t \in (-\pi,\pi]} \frac{\tilde{I}_{n,X}(\lambda_t)}{|\psi(e^{i\lambda_t},\beta)|^2},$$

where $\lambda_t = 2\pi t/n$, the Fourier frequencies. Let

$$\hat{\beta}_n := \arg\min_{\beta \in C} \tilde{\sigma}_n^2(\beta).$$

$\tilde{\beta}_n$ and $\hat{\beta}_n$ have the same asymptotic distributions:

$$\left(\frac{n}{\ln n}\right)^{1/\alpha} (\tilde{\beta}_n - \beta_0) \Rightarrow 4\pi W^{-1}(\beta_0) \sum_{j=1}^{\infty} b_j \frac{G_j}{G_0}$$

in $\mathbb{R}^{p+q}$ as $n \to \infty$, where $W^{-1}(\beta_0)$ is the inverse of the $(p+q) \times (p+q)$ matrix

$$W(\beta_0) = \int_{-\pi}^{\pi} \left[\frac{\partial \ln |\psi(e^{i\lambda}, \beta_0)|^2}{\partial \beta}\right] \left[\frac{\partial \ln |\psi(e^{i\lambda}, \beta_0)|^2}{\partial \beta}\right]' d\lambda,$$

$$b_j = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-ij\lambda} |\psi(e^{i\lambda}, \beta_0)|^2 \frac{\partial(|\psi(e^{i\lambda}, \beta_0)|^{-2})}{\partial \beta} d\lambda, \quad j \geq 1,$$

and G_0 and $\{G_j\}_{j\geq 1}$ are given in (4.4). For details, see Mikosch, et al. (1995).

Another possible approach to estimating β_0 is M-estimation, including the least absolute deviations (LAD) estimator. For an AR(p) model, the M-estimator $\hat{\phi}_n^{(M)}$ of $\phi = (\phi_1, \ldots, \phi_p)$ is

$$\hat{\phi}_n^{(M)} := \arg\min_{\phi} \rho(X_t - \phi_1 X_{t-1} - \ldots - \phi_p X_{t-p}),$$

where $\rho(\cdot)$ is a convex loss function satisfying certain conditions. Davis et al. (1992) show that

$$n^{1/\alpha}(\hat{\phi}_n^{(M)} - \phi_0) \Rightarrow \eta,$$

where η is a nondenerate random vector. This means that suitably constructed M-estimators are superior to the Whittle estimator in the sense of convergence rate. Davis (1996) shows that this is also true for heavy-tailed ARMA models.

Now consider the fractional ARIMA (FARIMA) model:

$$(1 - \phi_1 B - \cdots - \phi_p B^p)(1 - B)^d X_t = (1 + \theta_1 B + \cdots + \theta_q B^q) Z_t,$$

where $\{Z_t\}$ is a sequence of iid $S_\alpha(\sigma, 0, 0)$ rv's. Assume that the polynomials $\phi(z)$ and $\theta(z)$ satisfy the same assumptions as in ARMA models described before. Assume that $1 < \alpha < 2$ and $d \in (0, 1 - \frac{1}{\alpha})$. Let

$$\psi(z) := \frac{\theta(z)}{\phi(z)(1-z)^d} =: \sum_{j=0}^{\infty} \psi_j z^j, \quad |z| \leq 1.$$

Then it can be shown that $\psi_j = O(j^{d-1})$ as $j \to \infty$. So the process defined by the FARIMA model has long-range dependence (long-memory). Let

$$\beta = (\phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q, d)',$$

and β_0 be the true parameter. In this case, similar results hold for the Whittle estimator of β_0 as those in the case of ARMA models, though the arguments are more delicate. For details, see Kokoszka and Taqqu (1996).

7. Parameter estimation for regression models with heavy-tailed errors

We consider the regression model

$$Y_t = \beta_0 + \beta_1 x_t + X_t, \quad t \in \mathbb{N}.$$

For the sake of simplicity, we restrict to consider the case $x_t = t$, that is, the linear trend $\beta_0 + \beta_1 t$, although the results will hold under a rather general setting. We assume that $\{X_t\}$ is a sequence of iid $S_\alpha(\sigma, 0, 0)$ rv's. Our aim is to investigate asymptotic behaviors of the LSE and LAD estimator of the parameter. In a future study we want to deal with the case when $\{X_t\}$ is a heavy-tailed ARMA or FARIMA process, and also consider the asymptotic behavior of more general M-estimators.

Let $(\hat{\beta}_0, \hat{\beta}_1)$ denote the LSE of (β_0, β_1), which can be witten as

$$\hat{\beta}_0 = \beta_0 + \frac{\sum_{t=1}^{n}(\sum_{k=1}^{n} x_k^2 - x_t \sum_{k=1}^{n} x_k)X_t}{n\sum_{t=1}^{n} x_t^2 - (\sum_{t=1}^{n} x_t)^2} =: \beta_0 + \frac{\sum_{t=1}^{n} b_t X_t}{a_n},$$

$$\hat{\beta}_1 = \beta_1 + \frac{\sum_{t=1}^{n}(n x_t - \sum_{k=1}^{n} x_k)X_t}{n\sum_{t=1}^{n} x_t^2 - (\sum_{t=1}^{n} x_t)^2} =: \beta_1 + \frac{\sum_{t=1}^{n} c_t X_t}{a_n},$$

where $a_n = n\sum_{t=1}^{n} x_t^2 - (\sum_{t=1}^{n} x_t)^2$, $b_t = \sum_{k=1}^{n} x_k^2 - x_t \sum_{k=1}^{n} x_k$ and $c_t = n x_t - \sum_{k=1}^{n} x_k$. We have

THEOREM 1. *The LSE $(\hat{\beta}_0, \hat{\beta}_1)$ has the following distribution:*

$$n^{1-\frac{1}{\alpha}}(\hat{\beta}_0 - \beta_0) \sim S_\alpha(C_{n1}\sigma, 0, 0),$$
$$n^{2-\frac{1}{\alpha}}(\hat{\beta}_1 - \beta_1) \sim S_\alpha(C_{n2}\sigma, 0, 0),$$

where C_{n1} and C_{n2} are positive constants depending on n and α. Therefore, $\hat{\beta}_0$ is a consistent estimtor of β_0 if and only if $\alpha > 1$; $\hat{\beta}_1$ is a consistent estimtor of β_1 if and only if $\alpha > 1/2$.

Proof. We only consider $\hat{\beta}_0$, since the arguments for $\hat{\beta}_1$ are similar. $\sum_{t=1}^{n}(b_t/a_n)X_t$ has the distribution $S_\alpha((\sum_{t=1}^{n} |b_t|^\alpha)^{1/\alpha}/a_n, 0, 0)$. Now

$$\begin{aligned} a_n &= n\sum_{t=1}^{n} x_t^2 - (\sum_{t=1}^{n} x_t)^2 \\ &= n\sum_{t=1}^{n} t^2 - (\sum_{t=1}^{n} t)^2 \\ &= n^2(n+1)(2n+1)/6 - (n(n+1)/2)^2 \\ &= (n^4 - n^2)/12 \sim n^4/12 \end{aligned}$$

as $n \to \infty$.

$$
\begin{aligned}
\Big(\sum_{t=1}^{n} |b_t|^\alpha\Big)^{1/\alpha} &= \Big(\sum_{t=1}^{n}\Big|\sum_{k=1}^{n} k^2 - t\sum_{k=1}^{n} k\Big|^\alpha\Big)^{1/\alpha} \\
&= \Big(\sum_{t=1}^{n}\Big|\frac{n(n+1)(2n+1)}{6} - t\frac{n(n+1)}{2}\Big|^\alpha\Big)^{1/\alpha} \\
&= \frac{n(n+1)}{2}\Big(\sum_{t=1}^{n}\Big|\frac{2n+1}{3} - t\Big|^\alpha\Big)^{1/\alpha} \\
&= \frac{n(n+1)}{2}\Big(\sum_{t=1}^{[\frac{2n+1}{3}]}\Big|\frac{2n+1}{3} - t\Big|^\alpha + \sum_{[\frac{2n+1}{3}]+1}^{n}\Big|\frac{2n+1}{3} - t\Big|^\alpha\Big)^{1/\alpha} \\
&= \frac{n(n+1)}{2}\Big(\sum_{t=1}^{[\frac{2n-2}{3}]} t^\alpha + \sum_{t=1}^{[\frac{n-1}{3}]} t^\alpha\Big)^{1/\alpha} \\
&=: C'_{n1} n^2 n^{(1+\alpha)/\alpha} \\
&= C'_{n1} n^{3+\frac{1}{\alpha}}.
\end{aligned}
$$

Combining these results completes the proof. □

Next we consider the LAD estimator, which is defined as

$$(\tilde{\beta}_0, \tilde{\beta}_1) := \arg \min_{(u_0,u_1)\in\mathbb{R}^2} \sum_{t=1}^{n} |Y_t - u_0 - u_1 t|.$$

We have the following results.

THEOREM 2. *The LAD estimator $(\tilde{\beta}_0, \tilde{\beta}_1)$ has the following distribution:*

$$
\begin{aligned}
n^{1/2}(\tilde{\beta}_0 - \beta_0) &\Rightarrow \frac{2W_0 - 3W_1}{f(0)}, \\
n^{3/2}(\tilde{\beta}_1 - \beta_1) &\Rightarrow \frac{3(2W_1 - W_0)}{f(0)}
\end{aligned}
$$

for any $\alpha \in (0, 2]$, where $f(x)$ is the probability density of the regression error X_t, and (W_0, W_1) is normally distributed with mean $\mathbf{0}$ and covariance $\begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{3} \end{pmatrix}$.

REMARK 1. From Theorems 1 and 2, we see that the LAD estimator is largely superior to the LSE when the regression error has a heavy-tailed distribution.

Proof of Theorem 2. Theorem 2 will follow from Theorem 1 and Corollary 2 of Knight (1998), provided we can show that all the assumptions in Knight (1998) are satisfied in our case, which is not trivial at all. We note that in the proof of Theorem 1 in Knight (1998), many important details have not been given. Here we will give a detailed proof of our Theorem 2 using the idea of Knight (1998).

By the definition the LAD estimator, we have

$$\begin{aligned}(\tilde{\beta}_0, \tilde{\beta}_1) &:= \arg \min_{(u_0,u_1)\in\mathbb{R}^2} \sum_{t=1}^{n} |Y_t - u_0 - u_1 t| \\ &= \arg \min_{(u_0,u_1)\in\mathbb{R}^2} \sum_{t=1}^{n} (|Y_t - u_0 - u_1 t| - |X_t|) .\end{aligned}$$

Thus

$$\sum_{t=1}^{n} \left(|Y_t - \tilde{\beta}_0 - \tilde{\beta}_1 t| - |X_t|\right) = \min_{(u_0,u_1)\in\mathbb{R}^2} \sum_{t=1}^{n} (|Y_t - u_0 - u_1 t| - |X_t|) .$$

Note that

$$\begin{aligned}&\sum_{t=1}^{n} (|Y_t - u_0 - u_1 t| - |X_t|) \\ &= \sum_{t=1}^{n} (|X_t - (u_0 - \beta_0) - (u_1 - \beta_1)t| - |X_t|) \\ &= \sum_{t=1}^{n} \left(|X_t - n^{1/2}(u_0 - \beta_0)\frac{1}{n^{1/2}} - n^{3/2}(u_1 - \beta_1)\frac{t}{n^{3/2}}| - |X_t|\right),\end{aligned}$$

so let

$$Z_n(u_0, u_1) := \sum_{t=1}^{n} \left(\left|X_t - \frac{u_0}{n^{1/2}} - \frac{u_1 t}{n^{3/2}}\right| - |X_t|\right), \quad (u_0, u_1) \in \mathbb{R}^2,$$

then $(n^{1/2}(\tilde{\beta}_0 - \beta_0), n^{3/2}(\tilde{\beta}_1 - \beta_1))$ is the minimum point of the sample path of the stochastic process $\{Z_n(u_0, u_1)\}$ in $C(\mathbb{R}^2)$, which is the space of continuous functions on $\mathbb{R}^2$ equipped with the uniform metric. Therefore, if we can show that

$$Z_n(u_0, u_1) \Rightarrow Z(u_0, u_1) \quad \text{in } C(\mathbb{R}^2) \tag{7.1}$$

as $n \to \infty$, where $\Rightarrow$ denotes weak convergence and $\{Z(u_0, u_1)\}$ is a Gaussian process in $C(\mathbb{R}^2)$ whose sample path has an unique minimum point a.s. with the required normal distribution, then the proof will be completed.

Firstly let us show that (7.1) holds. Note that

$$|x - y| - |x| = -y[I(x > 0) - I(x < 0)] + 2 \int_0^y [I(x \le s) - I(x \le 0)]ds$$

holds for any $x \neq 0$, where $I(A)$ is the indicator function of the set A. From this, we can write

$$Z_n(u_0, u_1) = Z_n^{(1)}(u_0, u_1) + Z_n^{(2)}(u_0, u_1), \tag{7.2}$$

where

$$Z_n^{(1)}(u_0,u_1) := -\sum_{t=1}^{n}\left(\frac{u_0}{n^{1/2}}+\frac{u_1 t}{n^{3/2}}\right)[I(X_t>0)-I(X_t<0)] =: -\sum_{t=1}^{n} Z_{nt}^{(1)}(u_0,u_1) \tag{7.3}$$

and

$$Z_n^{(2)}(u_0,u_1) := 2\sum_{t=1}^{n}\int_0^{u_0/n^{1/2}+u_1 t/n^{3/2}}[I(X_t\le s)-I(X_t\le 0)]ds =: 2\sum_{t=1}^{n} Z_{nt}^{(2)}(u_0,u_1). \tag{7.4}$$

Now let us show that

$$Z_n^{(1)}(u_0,u_1) \Rightarrow -(u_0W_0+u_1W_1) \quad \text{in } C(\mathbb{R}^2) \tag{7.5}$$

as $n\to\infty$, where (W_0,W_1) is a 2-dimensional normally distributed random vector as given in Theorem 2. In order to show (7.5), it is enough to show that finite-dimensional distributions of $\{Z_n^{(1)}(u_0,u_1),\ \ (u_0,u_1)\in\mathbb{R}^2\}$ converge to those of $\{-(u_0W_0+u_1W_1),\ \ (u_0,u_1)\in\mathbb{R}^2\}$ and that the probability measures of $\{Z_n^{(1)}(u_0,u_1),$ $(u_0,u_1)\in\mathbb{R}^2, n\in\mathbb{N}\}$ are tight, by Theorems 8.1 and 8.2 of Billingsley (1968). We will only consider the univariate distributions, since multivariate case follows by using similar arguments and making use of the Cramér-Wold device.

Since the distribution of X_t is symmetric, we have $EZ_n^{(1)}(u_0,u_1)=0$. Furthermore,

$$\begin{aligned}
s_n^2 := \mathrm{Var}(Z_n^{(1)}(u_0,u_1)) &= \sum_{t=1}^{n}\left(\frac{u_0}{n^{1/2}}+\frac{u_1 t}{n^{3/2}}\right)^2 \mathrm{Var}(I(X_t>0)-I(X_t<0)) \\
&= \sum_{t=1}^{n}\left(\frac{u_0}{n^{1/2}}+\frac{u_1 t}{n^{3/2}}\right)^2 \\
&= u_0^2+(1+\frac{1}{n})u_0u_1+\frac{(n+1)(2n+1)}{6n^2}u_1^2 \\
&\to u_0^2+u_0u_1+\frac{u_1^2}{3}
\end{aligned} \tag{7.6}$$

as $n\to\infty$. We will show that

$$\frac{Z_n^{(1)}(u_0,u_1)}{s_n} \Rightarrow N(0,1) \tag{7.7}$$

as $n\to\infty$ for any $u_0\neq 0,\ u_1\neq 0$ (note that $Z_n^{(1)}(0,0)=0$ a.s.), which, in view of (7.6), is equivalent to

$$Z_n^{(1)}(u_0,u_1) \Rightarrow -(u_0W_0+u_1W_1), \quad \forall (u_0,u_1)\in\mathbb{R}^2\backslash(0,0),$$

as $n \to \infty$. To show (7.7), we will use the Lindeberg-Feller theorem (e.g. Shiryaev (1996); P.334). Write

$$\frac{Z_n^{(1)}(u_0, u_1)}{s_n} = -\sum_{t=1}^{n} \xi_{nt}, \quad n \in \mathbb{N},$$

where

$$\xi_{nt} := \frac{u_0/n^{1/2} + u_1 t/n^{3/2}}{s_n}(I(X_t > 0) - I(X_t < 0)), \quad 1 \le t \le n.$$

We will show that the triangular array $\{\xi_{nt} : n \in \mathbb{N}, 1 \le t \le n\}$ satisfies the conditions of the Lindeberg-Feller theorem. Clearly, $\{\xi_{n1}, \ldots, \xi_{nn}\}$ is independent with $E\xi_{nt} = 0$ and $\mathrm{Var}(\sum_{t=1}^{n} \xi_{nt}) = 1$. Also, we have

$$\begin{aligned}
E(\xi_{nt}^2) &= (u_0/n^{1/2} + u_1 t/n^{3/2})^2 E[(I(X_t > 0) - I(X_t < 0))^2]/s_n^2 \\
&= (u_0/n^{1/2} + u_1 t/n^{3/2})^2/s_n^2 \\
&= O(n^{-1}) \quad \text{uniformly in } t.
\end{aligned}$$

Therefore

$$\max_{1\le t\le n} E(\xi_{nt}^2) \to 0, \quad n \to \infty.$$

We have

$$\begin{aligned}
\sum_{t=1}^{n} E|\xi_{nt}|^4 &= \frac{1}{s_n^4}\sum_{t=1}^{n}\left(\frac{u_0}{n^{1/2}} + \frac{u_1 t}{n^{3/2}}\right)^4 E[(I(X_t > 0) - I(X_t < 0))^4] \\
&\le \frac{1}{s_n^4}\sum_{t=1}^{n}\left(\frac{|u_0|}{n^{1/2}} + \frac{|u_1|t}{n^{3/2}}\right)^4 \\
&= \frac{1}{s_n^4} O\left(\int_0^n \left(\frac{|u_0|}{n^{1/2}} + \frac{|u_1|t}{n^{3/2}}\right)^4 dt\right) \\
&= \frac{1}{s_n^4} O\left(\frac{n^{3/2}}{5|u_1|}\left[\left(\frac{|u_0|}{n^{1/2}} + \frac{|u_1|t}{n^{3/2}}\right)^5\right]_{t=0}^{n}\right) \\
&= O(1)O(n^{3/2}n^{-5/2}) \\
&= O(n^{-1})
\end{aligned}$$

as $n \to \infty$. This implies that the Lindeberg condition is satisfied, i.e.

$$\sum_{t=1}^{n} E[\xi_{nt}^2 I(|\xi_{nt}| \ge \varepsilon)] \to 0 \quad (n \to \infty; \forall \varepsilon > 0).$$

Combining the above results, we see that $\{\xi_{nt} : n \in \mathbb{N}, 1 \le t \le n\}$ satisfies the conditions of the Lindeberg-Feller theorem. Therefore we have proved (7.7).

Now we show that the probability measures of $\{Z_n^{(1)}(u_0, u_1), (u_0, u_1) \in \mathbb{R}^2, n \in \mathbb{N}\}$ are tight. Noticing that $Z_n^{(1)}(0,0) = 0$ a.s., in view of Theorem 8.2 of Billingsley (1968) it is sufficient to show that

$$\lim_{\delta\to 0} \limsup_{n\to\infty} P\left\{ \sup_{\|\mathbf{u}-\mathbf{v}\|\le\delta} |Z_n^{(1)}(\mathbf{u}) - Z_n^{(1)}(\mathbf{v})| > \eta \right\} = 0 \tag{7.8}$$

for any $\eta > 0$, where $\mathbf{u} = (u_0, u_1) \in \mathbb{R}^2$, $\mathbf{u} = (v_0, v_1) \in \mathbb{R}^2$ and $\|\cdot\|$ is the Euclidean norm on $\mathbb{R}^2$. First note that

$$\begin{aligned}
&\sup_{\|\mathbf{u}-\mathbf{v}\|\le\delta} |Z_n^{(1)}(\mathbf{u}) - Z_n^{(1)}(\mathbf{v})| \\
&= \sup_{\|\mathbf{u}-\mathbf{v}\|\le\delta} \left| \sum_{t=1}^{n} \left(\frac{u_0 - v_0}{n^{1/2}} + \frac{(u_1 - v_1)t}{n^{3/2}} \right) [I(X_t > 0) - I(X_t < 0)] \right| \\
&\le \sup_{\|\mathbf{u}-\mathbf{v}\|\le\delta} |u_0 - v_0| \frac{1}{n^{1/2}} \left| \sum_{t=1}^{n} [I(X_t > 0) - I(X_t < 0)] \right| \\
&+ \sup_{\|\mathbf{u}-\mathbf{v}\|\le\delta} |u_1 - v_1| \frac{1}{n^{3/2}} \left| \sum_{t=1}^{n} t[I(X_t > 0) - I(X_t < 0)] \right| \\
&\le \frac{\delta}{n^{1/2}} \left| \sum_{t=1}^{n} [I(X_t > 0) - I(X_t < 0)] \right| + \frac{\delta}{n^{3/2}} \left| \sum_{t=1}^{n} t[I(X_t > 0) - I(X_t < 0)] \right| \\
&\le \frac{\delta}{n^{1/2}} \left| \sum_{t=1}^{n} [I(X_t > 0) - I(X_t < 0)] \right| \\
&+ \frac{\delta}{n^{3/2}} 2n \max_{1\le k\le n} \left| \sum_{t=1}^{k} [I(X_t > 0) - I(X_t < 0)] \right|,
\end{aligned}$$

where the last inequality follows from making Abel's transformation of the sum in the second term (e.g. Zygmund (1993); P.3–4). Therefore, by Kolmogrov's maximal inequality (e.g. Shiryaev (1996); P.384) we have

$$\begin{aligned}
&P\left\{ \sup_{\|\mathbf{u}-\mathbf{v}\|\le\delta} |Z_n^{(1)}(\mathbf{u}) - Z_n^{(1)}(\mathbf{v})| > \eta \right\} \\
&\le P\left\{ \left| \sum_{t=1}^{n} [I(X_t > 0) - I(X_t < 0)] \right| > \frac{n^{1/2}\eta}{2\delta} \right\} \\
&+ P\left\{ \max_{1\le k\le n} \left| \sum_{t=1}^{k} [I(X_t > 0) - I(X_t < 0)] \right| > \frac{n^{1/2}\eta}{4\delta} \right\} \\
&\le \left(\frac{4\delta^2}{n\eta^2} + \frac{16\delta^2}{n\eta^2} \right) E\left[\sum_{t=1}^{n} (I(X_t > 0) - I(X_t < 0)) \right]^2 \\
&= \frac{20\delta^2}{\eta^2}
\end{aligned}$$

for each n, from which (7.8) follows. Now the proof of (7.5) is completed.

Now we consider $\{Z_n^{(2)}(\mathbf{u}), \mathbf{u} \in \mathbb{R}^2\}$. Recall that

$$Z_n^{(2)}(\mathbf{u}) := 2\sum_{t=1}^{n} \int_0^{u_0/n^{1/2}+u_1 t/n^{3/2}} (I(X_t \le s) - I(X_t \le 0))ds =: 2\sum_{t=1}^{n} Z_{nt}^{(2)}(\mathbf{u}).$$

Write

$$Z_n^{(2)}(\mathbf{u}) = 2\sum_{t=1}^{n} E(Z_{nt}^{(2)}(\mathbf{u})) + 2\sum_{t=1}^{n}[Z_{nt}^{(2)}(\mathbf{u}) - E(Z_{nt}^{(2)}(\mathbf{u}))]. \tag{7.9}$$

Since $X_1 \sim S_\alpha(\sigma, 0, 0)$ by assumption, X_1 has a continuous symmetric probability density. Let $F(x)$ and $f(x)$ denote the distribution function and the probability density of X_1, respectively. We have

$$\begin{aligned}
2\sum_{t=1}^{n} E(Z_{nt}^{(2)}(\mathbf{u})) &= 2\sum_{t=1}^{n} E\int_0^{u_0/n^{1/2}+u_1 t/n^{3/2}} (I(X_t \le s) - I(X_t \le 0))ds \\
&= 2\sum_{t=1}^{n} \int_0^{u_0/n^{1/2}+u_1 t/n^{3/2}} (F(s) - F(0))ds \\
&= \frac{2}{\sqrt{n}}\sum_{t=1}^{n} \int_0^{u_0+u_1 t/n} (F(s/\sqrt{n}) - F(0))ds \\
&= \frac{2}{n}\sum_{t=1}^{n} \int_0^{u_0+u_1 t/n} \sqrt{n}(F(s/\sqrt{n}) - F(0))ds \\
&=: \frac{2}{n}\sum_{t=1}^{n} \Psi_n(u_0 + u_1 t/n),
\end{aligned} \tag{7.10}$$

where

$$\Psi_n(t) := \int_0^t \sqrt{n}(F(s/\sqrt{n}) - F(0))ds.$$

Note that

$$\sqrt{n}(F(s/\sqrt{n}) - F(0)) \to sf(0), \quad n \to \infty,$$

uniformly for $s \in K$, where K is an arbitrary compact subset of $\mathbb{R}^1$, we have

$$\Psi_n(t) \to \int_0^t sf(0)ds = \frac{1}{2}f(0)t^2, \quad (n \to \infty, \forall t \in \mathbb{R}^1)$$

by Lebesgue's dominated convergence theorem, where, in particular, the convergence is uniform on arbitrary compact subset of $\mathbb{R}^1$. Thus, for any $\varepsilon > 0$, we can choose $n_0 = n_0(\varepsilon) \in \mathbf{N}$ such that

$$\left|\Psi_n(t) - \frac{1}{2}f(0)t^2\right| < \frac{\varepsilon}{2}, \quad (|t| \le |u_0| + |u_1|; n \ge n_0),$$

from which it holds that

$$\frac{2}{n}\sum_{t=1}^{n}\Psi_n(u_0+u_1t/n)=\frac{2}{n}\sum_{t=1}^{n}\frac{f(0)}{2}(u_0+u_1t/n)^2+\varepsilon,\quad (n\geq n_0), \tag{7.11}$$

where

$$\frac{2}{n}\sum_{t=1}^{n}\frac{f(0)}{2}(u_0+u_1t/n)^2\to f(0)\left(u_0^2+u_0u_1+\frac{u_1^2}{3}\right),\quad (n\to\infty). \tag{7.12}$$

Then (7.10), (7.11) and (7.12) show that

$$2\sum_{t=1}^{n}E(Z_{nt}^{(2)}(\mathbf{u}))\to f(0)\left(u_0^2+u_0u_1+\frac{u_1^2}{3}\right),\quad (n\to\infty). \tag{7.13}$$

Finally, we have

$$2\sum_{t=1}^{n}[Z_{nt}^{(2)}(\mathbf{u})-E(Z_{nt}^{(2)}(\mathbf{u}))]\xrightarrow{P}0,\quad (n\to\infty), \tag{7.14}$$

where $\xrightarrow{P}$ denotes convergence in probability, because (note that $Z_{nt}^{(2)}(\mathbf{u})\geq 0$ a.s.)

$$\begin{aligned}
&\mathrm{Var}\left(2\sum_{t=1}^{n}[Z_{nt}^{(2)}(\mathbf{u})-E(Z_{nt}^{(2)}(\mathbf{u}))]\right)\\
&=4\sum_{t=1}^{n}\mathrm{Var}(Z_{nt}^{(2)}(\mathbf{u}))\\
&\leq 4\sum_{t=1}^{n}E[(Z_{nt}^{(2)}(\mathbf{u}))^2]\\
&\leq 4\sum_{t=1}^{n}E\left[\int_0^{u_0/n^{1/2}+u_1t/n^{3/2}}(|u_0|/n^{1/2}+|u_1|t/n^{3/2})(I(X_t\leq s)-I(X_t\leq 0))ds\right]\\
&\leq 4\max_{1\leq t\leq n}(|u_0|/n^{1/2}+|u_1|t/n^{3/2})\sum_{t=1}^{n}E(Z_{nt}^{(2)}(\mathbf{u}))\\
&=O(n^{-1/2})
\end{aligned}$$

as $n\to\infty$ by (7.13). From (7.9), (7.13) and (7.14), it follows that

$$Z_n^{(2)}(\mathbf{u})\xrightarrow{P}f(0)\left(u_0^2+u_0u_1+\frac{u_1^2}{3}\right),\quad (n\to\infty). \tag{7.15}$$

From (7.5) and (7.15) we obtain

$$Z_n(\mathbf{u})\Rightarrow Z(\mathbf{u})\quad \text{in } C(\mathbb{R}^2) \tag{7.16}$$

as $n \to \infty$, where $\{Z(\mathbf{u}), \mathbf{u} \in \mathbb{R}^2\}$ is a Gaussian process in $C(\mathbb{R}^2)$ given by

$$Z(\mathbf{u}) = -(u_0 W_0 + u_1 W_1) + f(0)\left(u_0^2 + u_0 u_1 + \frac{u_1^2}{3}\right).$$

Note that the sample path of $\{Z(\mathbf{u}), \mathbf{u} \in \mathbb{R}^2\}$ has the unique minimum point a.s. which satisfies

$$\begin{cases} \dfrac{\partial Z(\mathbf{u})}{\partial u_0} = 2f(0)u_0 + f(0)u_1 - W_0 = 0, \\[2ex] \dfrac{\partial Z(\mathbf{u})}{\partial u_1} = f(0)u_0 + \dfrac{2f(0)}{3}u_1 - W_1 = 0, \end{cases}$$

whose solution is

$$\begin{cases} u_0 = \dfrac{2W_0 - 3W_1}{f(0)} \\[2ex] u_1 = \dfrac{3(2W_1 - W_0)}{f(0)}. \end{cases}$$

The proof of Theorem 2 is completed. □

REFFERENCES

[1] P.Billingsley, *Convergence of Probability Measures*, John Wiley & Sons, New York, 1968.

[2] R.A.Davis, *Gauss-Newton and M-estimation for ARMA processes*, Stochastic Process. Appl. **63** (1996), 75–95.

[3] R.A.Davis, K.Knight and J.Liu, *M-estimation for autoregressions with infinite variance*, Stochastic Process. Appl. **40** (1992), 145–180.

[4] R.A.Davis and S.I.Resnick, *Limit theory for moving averages of random variables with regularly varying tail probabilities*, Ann. Probab. **13** (1985), 179–195.

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

[6] W.Feller, *An Introduction to Probability Theory II*, Wiley, New York, 1971.

[7] T.Hida, *Brownian Motion*, Springer, Berlin, 1980.

[8] I.A.Ibragimov and Yu.V.Linnik, *Independent and Stationary Sequences of Random Variables*, Wolters–Noordhoff, Groningen, 1971.

[9] C.Klüppelberg and T.Mikosch, *Spectral estimates and stable processes*, Stochastic Process. Appl. **47** (1993), 323-344.

[10] C.Klüppelberg and T.Mikosch, *The integrated periodogram for stable processes*, Ann. Statist. **24** (1996), 1855–1879.

[11] K.Knight, *Limiting distributions for L_1 regression estimators under general conditions*, Ann. Statist. **26** (1998), 755–770.

[12] P.S.Kokoszka and M.S.Taqqu, *Parameter estimation for infinite variance fractional ARIMA*, Ann. Statist. **24** (1996), 1880–1913.

[13] T.Mikosch, T.Gadrich, C.Klüppelberg and R.Adler, *Parameter estimation for ARMA models with infinite variance innovations*, Ann. Statist. **23** (1995), 305–326.

[14] A.N.Shiryaev, *Probability* (2nd edn), Springer-Verlag, New York, 1996.

[15] A.Zygmund, *Trigonometric Series* (paperback edn), Vols.I,II, Cambridge University Press, Cambridge, 1993.

Department of Mathematical Sciences
Faculty of Science
University of the Ryukyus
Nishihara-cho, Okinawa 903-0213
JAPAN