应用概率统计 第二十四卷
第四期 2008年8月

Chinese Journal of Applied Probability
and Statistics Vol.24 No.4 Aug. 2008

Change-Point in the Mean of Heavy-Tailed Dependent Observations *

Han Sier[1] Tian Zheng[1,2] Wang Hongjun[1]

([1] *Department of Mathematics, Northwestern Polytechnical University, Xi'an, 710072*)

([2] *National Key Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing, 100080*)

Abstract

This paper studies the problem of mean change point in heavy-tailed dependent observations. We prove the consistency of CUSUM estimator of change-point and derive the rate of convergence. A Hájek-Rényi type inequality is also proved. Results are obtained under weak moment assumptions.

Keywords: Change-point estimation, heavy-tails, Hájek-Rényi inequality.

AMS Subject Classification: 60G52.

§1. Introduction

Estimation of the change point in the mean is often an important step in analysis of time series and has received considerable attention in the literature. Yao (1987)[1] considered the change point in a sequence of independent variables. Horváth (1997)[2] and many others estimated a shift in linear process. Some further references can be found in Kokoszka and Leipus (1998)[3] who focus on detecting change point in nonlinear time series. These authors investigate the processes which have finite variance.

In this paper we assume that the observations follow the model

$$X_t = \mu(t) + Y_t, \tag{1.1}$$

where $\mu(t)$ is a nonstochastic function in time and Y_t is a zero-mean stationary time series with heavy-tailed unvariate marginal distributions. We assume that these distributions regularly vary with index κ satisfying $1 < \kappa < 2$, so that the mean exists but the variance is infinite which is difference from the aforementioned work.

*This research is supported by the National Natural Science Foundation of China (No. 60375003).
Received September 27, 2004.

We consider the simple case where $\mu(t)$ only takes two difference values, μ_1 before time k^* and μ_2 after time k^*. That is,

$$\mu(t) = \begin{cases} \mu_1, & \text{if } \ t \le k^*; \\ \mu_2, & \text{if } \ t > k^*, \end{cases} \tag{1.2}$$

where μ_1, μ_2 and k^* are unknown and k^* is a change point.

The goal of this paper is to estimate k^* given n observations $X_1, X_2, \cdots, X_n$. The CUSUM-type estimator $\widehat{k}^*$ of the change point k^* in the mean is defined as follows:

$$\widehat{k}^* = \min\Big\{k : |U_k| = \max_{1\le j<n} |U_j|\Big\}, \tag{1.3}$$

where

$$U_k = \Big(\frac{k(n-k)}{n^{1+\beta}}\Big)^{1-\gamma}\Big\{\frac{1}{k}\sum_{j=1}^{k} X_j - \frac{1}{n-k}\sum_{j=k+1}^{n} X_j\Big\} \tag{1.4}$$

with some $0 \le \gamma < 1$ and $\beta > 0$.

In this paper, we extend the theory of Kokoszka and Leipus (1998)[3], which studies the the CUSUM-type estimators of the change in the mean of dependent observations with finite variance, to processes with infinite-variance heavy-tailed distributions.

§2. Assumptions

In this paper, we assume that the time series under consideration is strong mixing and satisfy the following Assumptions:

Assumption 2.1 The sequence Y_t is stationary with symmetric univariate marginal distributions which satisfy

$$n\mathsf{P}(Y_1/a_n \in \cdot) \xrightarrow{v} \mu(\cdot) \tag{2.1}$$

with the a_n defined by $n\mathsf{P}(|Y_1| > a_n) \longrightarrow 1$ and the measure μ given by

$$2\mu(\mathrm{d}x) = \kappa|x|^{-\kappa-1}\mathbf{1}\{x<0\}\mathrm{d}x + \kappa x^{-\kappa-1}\mathbf{1}\{x>0\}\mathrm{d}x \tag{2.2}$$

and $\xrightarrow{v}$ denotes vague convergence on $R-\{0\}$. Moreover we assume that

$$\sum_{t=1}^{n} \delta_{Y_t/a_n} \xrightarrow{d} \sum_{i=1}^{\infty}\sum_{i=1}^{\infty} \delta_{P_i Q_{ij}} \tag{2.3}$$

with the limiting point process as in Theorem 2.3 and Corollary 2.4 of Davis and Hsing (1995)[4].

Assumption 2.2 For every $1 \le \delta < \kappa$, the observations $\{X_k, 1 \le k \le n\}$ defined by (2.1) satisfy

$$\max_{1\le k\le n} \mathsf{E}(X_k - \mathsf{E}X_k)^{\delta} \le C. \tag{2.4}$$

Throughout the paper C stands for a positive constant whose value dose not depend on n and may vary from formula to formula.

Condition (2.1) together with (2.2) is said to be conditions of "heavy tails" and equivalent to the requirement that the one-dimensional marginal distributions are in the domain of attraction of a κ-stable law. Condition (2.3) is implied by a very weak form of mixing assumed by Davis and Hsing (1995)[4], which in turn is implied by strong mixing conditions. If we assume, as we do in this paper, that a stochastic process has infinite variance, we cannot apply the Hájek-Rényi inequality established in Kokoszka and Leipus (2000)[3] where the stochastic process has finite variance. Instead we assume condition (2.4) under which the generalized Hájek and Rényi inequality in this paper is valid.

§3. Main Results

Theorem 3.1 Consider the sample $X_1, X_2, \cdots, X_n$ from model (1.1) and the change-point estimator $\widehat{k}^*$ given by (1.3). If Assumption 2.1 and 2.2 hold, then for $\widehat{\tau}^* = \widehat{k}^*/n^{(1+\beta)}$ we have

$$\mathsf{P}\{|\widehat{\tau}^* - \tau^*| > \epsilon\} \le \frac{C}{|\Delta|^\delta \epsilon^\delta} n^{\delta(\gamma\beta - 2\beta - 1/2)+1}, \tag{3.1}$$

where $\Delta = \mu_1 - \mu_2$, γ, β and δ satisfy $\gamma\beta - 2\beta < -1/2$ and $\delta < \kappa$.

The proof of the Theorem 3.1 is based on the Theorem 3.2 which is stated below.

Let ε_i's are i.i.d. zero-mean random variables with $\mathsf{E}\varepsilon_i^2 = \sigma^2$ and $\{c_k\}$ be a decreasing sequence of constants. Hájek and Rényi (1955)[5] proved that

$$\mathsf{P}\Big\{\max_{m\le k\le n} c_k \Big|\sum_{i=1}^k \varepsilon_i\Big| > \epsilon\Big\} \le \frac{\sigma^2}{\epsilon^2}\Big(mc_m^2 + \sum_{i=m+1}^n c_i^2\Big). \tag{3.2}$$

This inequality was later generalized to any random variables with finite variance by Kokoszka and Leipus (1998)[3]. We now generalize this inequality to heavy-tailed process with infinite variance.

Theorem 3.2 Let $X_1, X_2, cdots, X_n$ be defined by (2.1) and $1 < \delta < \kappa < 2$, the generalized Hájek and Rényi inequality takes the following form:

$$\begin{aligned} &\epsilon^\delta \mathsf{P}\Big\{\max_{m\le k\le n} c_k \Big|\sum_{i=1}^k X_i\Big| > \epsilon\Big\} \\ \le\ & c_m^\delta \mathsf{E}\Big|\sum_{j=1}^m X_j\Big|^\delta + \sum_{k=m+1}^{n-1}\Big\{|c_{k+1}^\delta - c_k^\delta|\mathsf{E}\Big|\sum_{j=1}^k X_j\Big|^\delta \\ &+ 2^\delta c_{k+1}^\delta\Big(\mathsf{E}\Big|\sum_{j=1}^k X_j\Big|^\delta\Big)^{1/2}(\mathsf{E}|X_{k+1}|^\delta)^{1/2} + c_{k+1}^\delta \mathsf{E}|X_{k+1}|^\delta\Big\}. \end{aligned} \tag{3.3}$$

Proof Using Theorem 3.1 of Kokoszka and Leipus (1998)[3], we get that for any random variables $M_1,\cdots,M_n$ and events $A=\Big\{\max\limits_{1\le k\le n} M_k>\epsilon\Big\}$, $D_k=\{M_1\le\epsilon,\cdots,M_k\le\epsilon\}$,

$$\epsilon 1_A \le M_1+\sum_{k=1}^{n-1}(M_{k+1}-M_k)1_{D_k}-M_n 1_{D_n}. \tag{3.4}$$

Let $M_k=c_k^\delta\Big|\sum\limits_{j=1}^k X_j\Big|^\delta$. Utilizing (3.4), we get

$$\begin{aligned}&\epsilon^\delta \mathsf{P}\Big\{\max_{m\le k\le n} c_k^\delta\Big|\sum_{j=1}^k X_j\Big|^\delta>\epsilon^\delta\Big\}\\ \le\ & c_m^\delta\mathsf{E}\Big|\sum_{j=1}^m X_j\Big|^\delta+\mathsf{E}\sum_{k=m}^{n-1}\Big(c_{k+1}^\delta\Big|\sum_{j=1}^{k+1}X_j\Big|^\delta-c_k^\delta\Big|\sum_{j=1}^k X_j\Big|^\delta\Big)1_{D_k}.\end{aligned} \tag{3.5}$$

By the c_r inequality and the Hölder inequality we have

$$\begin{aligned}&\mathsf{E}\Big(c_{k+1}^\delta\Big|\sum_{j=1}^{k+1}X_j\Big|^\delta-c_k^\delta\Big|\sum_{j=1}^k X_j\Big|^\delta\Big)1_{D_k}\\ =\ &\mathsf{E}\Big\{c_{k+1}^\delta\Big(\Big|\sum_{j=1}^k X_j+X_{j+1}\Big|^2\Big)^{\delta/2}-c_k^\delta\Big|\sum_{j=1}^k X_j\Big|^\delta\Big\}1_{D_k}\\ \le\ &\mathsf{E}\Big\{\Big(c_{k+1}^2\Big|\sum_{j=1}^k X_j\Big|^2+2c_{k+1}^2\Big|\sum_{j=1}^k X_j\Big||X_{k+1}|+c_{k+1}^2X_{k+1}^2\Big)^{\delta/2}-c_k^\delta\Big|\sum_{j=1}^k X_j\Big|^\delta\Big\}1_{D_k}\\ \le\ &\mathsf{E}\Big\{|c_{k+1}^\delta-c_k^\delta|\Big|\sum_{j=1}^k X_j\Big|^\delta+2^\delta c_{k+1}^\delta\Big(\Big|\sum_{j=1}^k X_j\Big||X_{k+1}|\Big)^{\delta/2}+c_{k+1}^\delta X_{k+1}^\delta\Big\}1_{D_k}\\ \le\ &|c_{k+1}^\delta-c_k^\delta|\mathsf{E}\Big|\sum_{j=1}^k X_j\Big|^\delta+2^\delta c_{k+1}^\delta\Big(\mathsf{E}\Big|\sum_{j=1}^k X_j\Big|^\delta\Big)^{1/2}(\mathsf{E}|X_{k+1}|^\delta)^{1/2}+c_{k+1}^\delta\mathsf{E}|X_{k+1}|^\delta,\end{aligned} \tag{3.6}$$

which yields (3.3). □

The proof of the Theorem 3.1.

Proof Let $\tau=k/n^{(1+\beta)}$, we have

$$\mathsf{E}U_k=\begin{cases}\Delta n^{(1-\gamma)(1+\beta)}\tau^{1-\gamma}(n^{-\beta}-\tau^*)(n^{-\beta}-\tau)^{-\gamma}, & \text{if}\ \ k\le k^*\\ \Delta n^{(1-\gamma)(1+\beta)}(n^{-\beta}-\tau)^{1-\gamma}\tau^*\tau^{-\gamma}, & \text{if}\ \ k>k^*\end{cases} \tag{3.7}$$

and

$$\mathsf{E}U_{k^*}=\Delta n^{(1-\gamma)(1+\beta)}(\tau^*)^{1-\gamma}(n^{-\beta}-\tau^*)^{1-\gamma}. \tag{3.8}$$

If $k\le k^*$, we have

$$\begin{aligned}|\mathsf{E}U_{k^*}|-|\mathsf{E}U_k| &= |\Delta|n^{(1-\gamma)(1+\beta)}(n^{-\beta}-\tau^*)^{1-\gamma}\Big((\tau^*)^{1-\gamma}-\tau^{1-\gamma}\Big(\frac{n^{-\beta}-\tau^*}{n^{-\beta}-\tau}\Big)^\gamma\Big)\\ &\ge |\Delta|n^{(1-\gamma)(1+\beta)}(n^{-\beta}-\tau^*)^{1-\gamma}((\tau^*)^{1-\gamma}-\tau^{1-\gamma}).\end{aligned} \tag{3.9}$$

By the mean value theorem,

$$(\tau^*)^{1-\gamma} - \tau^{1-\gamma} \geq (1-\gamma)(\tau^*)^{-\gamma}(\tau^* - \tau).$$

Thus we get

$$|\mathsf{E}U_{k^*}| - |\mathsf{E}U_k| \geq |\Delta| n^{(1-\gamma)(1+\beta)}(n^{-\beta} - \tau^*)^{1-\gamma}(1-\gamma)(\tau^*)^{-\gamma}(\tau^* - \tau). \tag{3.10}$$

Similarly, if $k \geq k^*$

$$\begin{aligned} |\mathsf{E}U_{k^*}| - |\mathsf{E}U_k| &= |\Delta| n^{(1-\gamma)(1+\beta)}(\tau^*)^{1-\gamma}\Big((n^{-\beta} - \tau^*)^{1-\gamma} - (n^{-\beta} - \tau)^{1-\gamma}\Big(\frac{\tau^*}{\tau}\Big)^{\gamma}\Big) \\ &\geq |\Delta| n^{(1-\gamma)(1+\beta)}(\tau^*)^{1-\gamma}((n^{-\beta} - \tau^*)^{1-\gamma} - (n^{-\beta} - \tau)^{1-\gamma}) \\ &\geq |\Delta| n^{(1-\gamma)(1+\beta)}(\tau^*)^{1-\gamma}(1-\gamma)(n^{-\beta} - \tau^*)^{-\gamma}(\tau - \tau^*). \end{aligned} \tag{3.11}$$

Combining (3.10) and (3.11), we obtained

$$|\mathsf{E}U_{k^*}| - |\mathsf{E}U_k| \geq |\Delta| n^{(1-\gamma)(1+\beta)}\overline{\tau}|\tau^* - \tau|, \tag{3.12}$$

where $\overline{\tau} := (1-\gamma)(\tau^*)^{-\gamma}(n^{-\beta} - \tau^*)^{-\gamma}\min\{\tau^*, n^{-\beta} - \tau^*\}$.

Observe that

$$\begin{aligned} |U_k| - |U_{k^*}| &\leq |U_k - \mathsf{E}U_k| + |\mathsf{E}U_k| + |U_{k^*} - \mathsf{E}U_{k^*}| - |\mathsf{E}U_{k^*}| \\ &\leq 2\max_{1\leq k\leq n}|U_k - \mathsf{E}U_k| + |\mathsf{E}U_k| - |\mathsf{E}U_{k^*}|. \end{aligned} \tag{3.13}$$

we get from (3.12)

$$|\Delta| n^{(1-\gamma)(1+\beta)}\overline{\tau}|\tau^* - \tau| \leq |\mathsf{E}U_{k^*}| - |\mathsf{E}U_k| \leq 2\max_{1\leq k\leq n}|U_k - \mathsf{E}U_k| + |U_{k^*}| - |U_k|. \tag{3.14}$$

Replacing τ by $\widehat{\tau}^*$ and noting that $|U_{k^*}| \leq |U_{\widehat{k}^*}|$, we get

$$\begin{aligned} |\Delta|\overline{\tau}|\tau^* - \tau| &\leq 2n^{(\gamma-1)(1+\beta)}\max_{1\leq k\leq n}|U_k - \mathsf{E}U_k| \\ &\leq 2n^{(\gamma-1)(1+\beta)-\beta}\Big\{\max_{1\leq k\leq n}\frac{1}{k^{\gamma}}\Big|\sum_{j=1}^{k}Y_j\Big| + \max_{1\leq k\leq n}\frac{1}{(n-k)^{\gamma}}\Big|\sum_{j=k+1}^{n}Y_j\Big|\Big\}. \end{aligned} \tag{3.15}$$

Applying Theorem 3.2 with $c_k = n^{(\gamma-1)(1+\beta)-\beta}k^{-\gamma}$, we have that

$$\begin{aligned} &\epsilon^{\delta}\mathsf{P}\Big\{n^{(\gamma-1)(1+\beta)-\beta}\max_{1\leq k\leq n}\frac{1}{k^{\gamma}}\Big|\sum_{j=1}^{k}Y_j\Big| > \epsilon\Big\} \\ \leq\ & n^{\delta\{(\gamma-1)(1+\beta)-\beta\}}\Big[\sum_{k=1}^{n-1}\Big(\frac{1}{k^{\delta\gamma}} - \frac{1}{(k+1)^{\delta\gamma}}\Big)\mathsf{E}\Big|\sum_{j=1}^{k}Y_j\Big|^{\delta} \\ &+ 2^{\delta}\sum_{k=1}^{n-1}\frac{1}{(k+1)^{\delta\gamma}}\Big(\mathsf{E}\Big|\sum_{j=1}^{k}Y_j\Big|^{\delta}\Big)^{1/2}(\mathsf{E}|Y_{k+1}|^{\delta})^{1/2} + \sum_{k=1}^{n-1}\frac{1}{(k+1)^{\delta\gamma}}\mathsf{E}|Y_{k+1}|^{\delta}\Big]. \end{aligned} \tag{3.16}$$

By Assumption 2.1 and Minkowski inequality, we get that

$$\mathsf{E}\Big|\sum_{j=1}^{k} Y_j\Big|^{\delta} \le \Big(\sum_{j=1}^{k}(\mathsf{E}|Y_j|^{\delta})^{1/\delta}\Big)^{\delta} \le Ck^{\delta}$$

and using the inequality

$$\frac{1}{k^{\delta\gamma}} - \frac{1}{(k+1)^{\delta\gamma}} \le \frac{\delta\gamma}{k^{\delta\gamma+1}},$$

we obtain from (3.16)

$$\begin{aligned}
&\epsilon^{\delta}\mathsf{P}\Big\{n^{(\gamma-1)(1+\beta)-\beta}\max_{1\le k\le n}\frac{1}{k^{\gamma}}\Big|\sum_{j=1}^{k}Y_j\Big|>\epsilon\Big\}\\
\le\;& Cn^{\delta\{(\gamma-1)(1+\beta)-\beta\}}\Big[\sum_{k=1}^{n-1}k^{\delta-\delta\gamma-1}\\
&+2^{\delta}\max_k(\mathsf{E}|Y_k|^{\delta})^{1/2}\sum_{k=1}^{n-1}k^{\delta/2-\delta\gamma}+\max_k(\mathsf{E}|Y_k|^{\delta})\sum_{k=1}^{n-1}(k+1)^{-\delta\gamma}\Big]\\
\le\;& Cn^{\delta\{(\gamma-1)(1+\beta)-\beta\}}\sum_{k=1}^{n}k^{\delta/2-\delta\gamma}\le Cn^{\delta(\gamma\beta-2\beta-1/2)+1}.
\end{aligned}\tag{3.17}$$

Similarly argument to (3.17), we can obtain

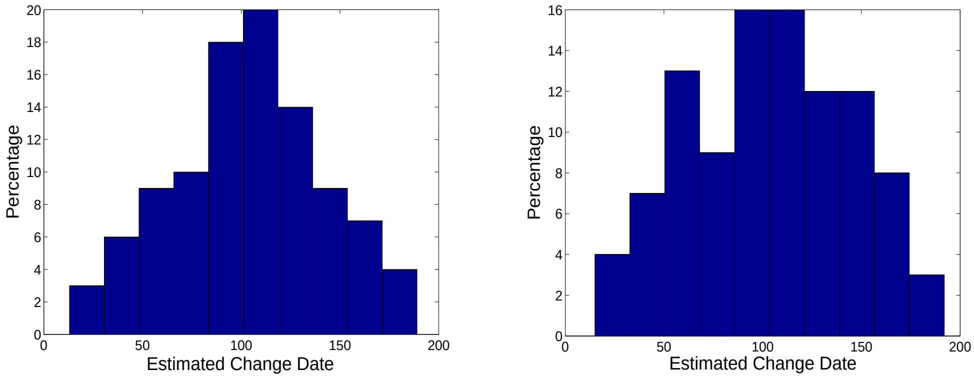
$$\epsilon^{\delta}\mathsf{P}\Big\{n^{(\gamma-1)(1+\beta)-\beta}\max_{1\le k\le n}\frac{1}{(n-k)^{\gamma}}\Big|\sum_{j=k+1}^{n}Y_j\Big|>\epsilon\Big\}\le Cn^{\delta(\gamma\beta-2\beta-1/2)+1}.\tag{3.18}$$

Combining (3.17) and (3.18) with (3.15) yields (3.1). □

§4. Simulation Study

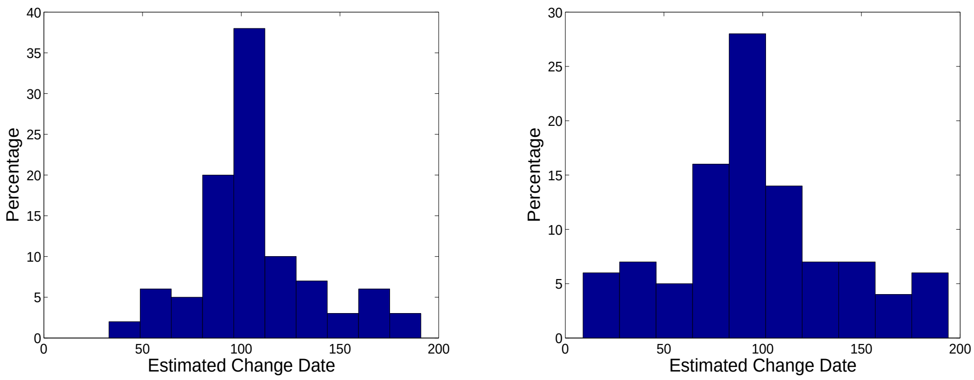
In this section, we briefly discuss the sample distribution of the change-point estimator $\widehat{k}^*$.

We consider two Data Generating Processes (DGP) as in Kokoszka and Wolf (2004)[6], who established the validity of subsampling confidence intervals for mean of a dependent series with heavy-tailed marginal distributions. The first DGP is an AR(1) model with stable innovations. That is, $X_t = \mu(t) + Y_t$, where $Y_t = \phi Y_{t-1} + Z_t$ and Z_t's are stable innovations with index κ. The second DGP is a GARCH(1,1) model: $X_t = \mu(t) + Y_t$, where $Y_t = \sigma_t\varepsilon_t$, $\sigma_t^2 = \omega + \alpha_1 Y_{t-j}^2 + \beta_1\sigma_{t-j}^2$ and ε_t's are standard normal innovations. By choosing positive values for ω, α_1 and β_1, such that the equation $\mathsf{E}(\alpha_1\varepsilon_1^2+\beta_1)^{\kappa/2} = 1$ has a solution $1 < \kappa < 2$, we can generate GARCH(1,1) time series with finite mean but infinite variance.



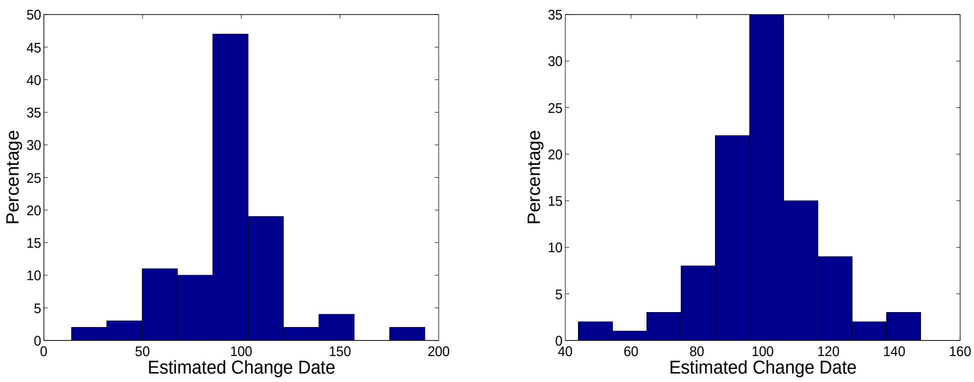
(a)　　(b)

Figure 1　Sampling distribution of $\widehat{k}^*$, $\kappa = 1.19$,
(a) $\omega = 1$, $\alpha_1 = 1.3$, $\beta_1 = 0.05$. (b) $\phi = 0.5$.



(a)　　(b)

Figure 2　Sampling distribution of $\widehat{k}^*$, $\kappa = 1.43$,
(a) $\omega = 1$, $\alpha_1 = 1.1$, $\beta_1 = 0.1$. (b) $\phi = 0.5$.



(a)　　(b)

Figure 3　Sampling distribution of $\widehat{k}^*$, $\kappa = 1.83$,
(a) $\omega = 1$, $\alpha_1 = 0.9$, $\beta_1 = 0.15$. (b) $\phi = 0.5$.

Experiments are carried out for $n=200$, $k^*=100$, $\mu_1=1$, $\mu_2=2$ and $\phi=0.5$. Figures 1-3 plot the histograms of 100 replications of $\widehat{k}^*$ for various combination of parameters, where Figure 1(a), 2(a) and 3(a) describe the sample distribution of $\widehat{k}^*$ for GARCH(1,1) models while Figure 1(b), 2(b) and 3(b) for AR(1) models. Obviously, the accuracy increases as κ gets closer to 2. For fixed κ, the range of $\widehat{k}^*$ for GARCH(1,1) models is smaller than that for AR(1) models, which shows that the estimated results of change point for GARCH(1,1) models is better than that for AR(1) models. Interestingly, similar findings had been obtained by Kokoszka and Wolf (2004)[6] who interpreted these phenomena as that there exist many "outliers" in observations when κ approaches 1 and realizations of GARCH(1,1) models do not exhibit isolated spikes but rather "clusters of high volatility".

Acknowledgements We would like to thank professor John Nolan who provide the software for generating the stable innovations used in Section 4.

References

[1] Yao, Y.C., Approximating the distribution of the ML estimate of the change-point in a sequence of independent r.v.'s, *Ann. Statist.*, **3(6)**(1987), 1321–1328.

[2] Horváth, L., *Limit Theorems in Change-Point Analysis*, Wiley, Chichester, 1997.

[3] Kokoszka, P. and Leipus, R., Change-point in the mean of dependent observations, *Statistics and Probability Letters*, **40(4)**(1998), 385–393.

[4] Davis, R.A. and Hsing, T., Point process and partial sum convergence for weakly dependent random variables with infinite variance, *The Annals of Probability*, **23(2)**(1995), 879 - 917.

[5] Hájek, J., Rényi, A., Generalization of an inequality of Kolmogorov, *Acta. Math. Acad. Sci. Hungar.*, **6(3)**(1955), 281–283.

[6] Kokoszka, P. and Wolf, M., Subsampling the mean of heavy-tailed dependent observations, *Journal of time series analysis*, **25(2)**(2004), 217–234.

厚尾相依序列的均值变点估计

韩四儿[1] 田 铮[1,2] 王红军[1]

([1]西北工业大学应用数学系, 西安, 710072)

([2]模式识别国家重点实验室, 中国科学院自动化研究所, 北京, 100080)

本文研究了厚尾相依序列的均值变点估计. 证明了变点的CUSUM估计的一致性并得到了收敛速度. 在方差无穷的情况下推广了Hájek–Rényi不等式.

关键词: 变点估计, 厚尾, Hájek-Rényi不等式.

学科分类号: O212.2.