应用概率统计 第二十九卷
第五期 2013年10月

Chinese Journal of Applied Probability
and Statistics Vol.29 No.5 Oct. 2013

独立随机序列均值多变点的非参数检测 *

秦瑞兵[1,2] 田 铮[2] 陈占寿[2,3]

([1]山西大学数学科学学院, 太原, 030006; [2]西北工业大学应用数学系, 西安, 710072)

([3]青海师范大学数学系, 西宁, 810008)

摘 要

本文考虑独立随机序列均值多变点的检测与估计问题, 提出一种非参数的检验方法, 给出其渐近分布, 同时亦给出了变点位置的估计, 并证明了变点个数估计的一致性. Monte Carlo试验研究了该检验统计量的有限样本性质, 结果表明该统计量对于厚尾误差具有较好势和经验水平. 最后将文中所给的非参数检验方法应用到鲁北化工(LBC)股票价格数据中, 结果表明, 文中所给统计量可以准确检验出多变点并进行估计.

关键词: 独立随机序列, 多变点, 非参数检验, 样本中位数.

学科分类号: O212.1.

§1. 引 言

变点问题最早起源于质量控制工程. 1954年, Page (1954)发表了第一篇关于变点统计分析的文章, 他研究自动生产线上产品质量检测问题. 当产品质量超过产品质量控制警界线时, 希望能及时预报, 以免生产出更多的次品. 这个质量发生质变的时刻, 即为变点. 变点不仅在工业自动化领域有广泛应用, 在经济、金融、气象、地质, 以及环保、公共卫生等各种领域均有广泛应用. 因此变点统计问题吸引了众多统计学家的关注. 近20年来, 变点的理论分析[2]与实际应用[3–7]都得到了充分的研究. 由于在实际应用中, 数据可能含有多个变点, 因此有必要检验多变点的存在, 并对变点的个数以及位置做出估计. 对于多变点的检测与估计, 目前较好的结果有Bai (1997, 1999)关于线性相依序列的多变点问题的工作, Lavielle和Moulines (2000)对线性序列均值多变点问题的工作, 韩四儿等(2006)对条件异方差自回归(ARCH)系数多变点问题的研究.

另一方面, 为了分析的方便, 现有理论大多建立在Gauss误差假设的基础上. 而实际应用领域, 例如通讯[12], 遥感数据处理[13], 更多的误差是非Gauss的, 甚至是厚尾的. 因此, 近年来, 对于厚尾序列的研究亦是方兴未艾, 特别是方差无穷情形的相关理论. 对于独立的方差无穷序列变点的研究, 最近也得到一些深刻的结论, 如张晨等(2007)以及Shi等(2009)关

*自然科学基金数学天元青年基金(11226217)、自然科学基金青年基金项目(11301291)、第51批中国博士后科学基金面上资助(2012M510772)和中国博士后科学基金第六批特别资助项目(2013T60266)资助.

本文2009年11月11日收到, 2013年7月18日收到修改稿.

于均值变点的工作. 这些统计量由于要求一阶矩或二阶矩有穷, 不能直接应用到诸如柯西分布的更加厚尾的新息过程的均值变点问题中.

本文考虑独立随机序列的均值多变点检测与估计问题, 构造了一种非参数的检验统计量, 并得到与通常所用的CUSUM统计量相同的渐近结果. 模拟表明, 对于厚尾序列, 本文的检验方法具有较高的势和较好的经验水平. 最后, 将文中方法应用到金融数据的均值变点检测之中, 结果表明, 文中所给统计量可以准确检验出多变点并进行估计.

§2. 模型与假设

考虑如下均值变点模型

$$Y_t=\begin{cases}\mu_1+X_t, & t=1,2,\ldots,k_1^0;\\ \mu_2+X_t, & t=k_1^0+1,\ldots,k_2^0;\\ \quad\vdots & \\ \mu_{m+1}+X_t, & t=k_m^0+1,\ldots,T.\end{cases}\tag{2.1}$$

其中X_t满足如下假设:

假设 2.1 1. $\{X_t,1\le t\le T\}$为独立同分布序列, 分布函数为$F(\cdot)$且具有相同的中位数$\widetilde{\mu}=0$;

2. 对任意$\eta>0$, 使得$X_t-\widetilde{\mu}$的密度函数$f(\cdot)$在区间$[-\eta,\eta]$上连续, 且有$\inf\limits_{x\in[-\eta,\eta]}f(x)>0$.

假设 2.2 存在$\tau_i^0\in(0,1)$, $\eta>0$, 使得$k_i^0=[T\tau_i^0]$, 并且$\tau_{i+1}^0-\tau_i^0\ge\eta$, $i=0,1,\ldots,m$, 其中$[\cdot]$为取整函数.

注记 1 假设2.1对X_t无矩要求, 故X_t可以包括柯西分布以及尾部更厚的序列. 假设2.2确保模型(2.1)的每两个变点之间有足够多的样本, 这是大数定理和泛函中心极限定理成立的基本条件, 通常η取0.05, 0.01等较小的数.

为讨论方便, 首先给出一些记号: $[i,j]$表示由时刻i到时刻j观测值的集合, $\Delta_{[i,j]}$表示区间$[i,j]$上的均值跳跃度, $\xrightarrow{\mathrm{P}}$表示依概率收敛, $\xrightarrow{\mathrm{d}}$表示依分布收敛, $\xrightarrow{\mathrm{a.s.}}$表示几乎处处收敛, 当$i\le 0$时, $k_i^0=0$, 当$i\ge m$时, $k_i^0=T$.

已知至少有l个变点$k_1^0,k_2^0,\ldots,k_l^0$, 若需要估计第$l+1$个变点, 则构造统计量

$$\begin{aligned}(\widehat{k}_1^0,\ldots,\widehat{k}_{l+1}^0) &= \arg\max_{k_1,\ldots,k_{l+1}} R(k_1,\ldots,k_{l+1})\\ &= \arg\max_{k_1,\ldots,k_{l+1}}\sum_{i=1}^{l+1}R_{[k_{i-1},k_{i+1}]}(k_i),\end{aligned}\tag{2.2}$$

其中$k_1,\ldots,k_{l+1}$中有且仅有l个与$k_1^0,\ldots,k_l^0$相同,

$$\begin{aligned} R_{[k_{i-1},k_{i+1}]}(k_i) &= \frac{(k_i-k_{i-1})(k_{i+1}-k_i)}{(k_{i+1}-k_{i-1})^2}\times\Big|\frac{1}{k_i-k_{i-1}}\sum_{t=k_{i-1}+1}^{k_i}\mathrm{sgn}(Y_t-m_T) \\ &\quad -\frac{1}{k_{i+1}-k_i}\sum_{t=k_i+1}^{k_{i+1}}\mathrm{sgn}(Y_t-m_T)\Big|, \end{aligned} \tag{2.3}$$

其中m_T为样本$X_1,\ldots,X_T$的样本中位数.

第$l+1$个变点的非参数型CUSUM估计定义为

$$\widehat{k}_{l+1}^* = \{\widehat{k}_i^0 : \widehat{k}_i^0 \neq k_j^0, i=1,\ldots,l=1; j=1,\ldots,l\}.$$

令$\tau_i^0=k_i^0/T$, $\widehat{\tau}_i^0=\widehat{k}_i^0/T$, 若$|\widehat{\tau}_i^0-\tau_i^0|=o_P(1)$, 则称$\widehat{k}_i^0$为$k_i^0$的$T$相合估计, $i=1,\ldots,m$.

§3.　主要结果

本节考虑原假设$H_0: m=l$, 备择假设$H_1: m=l+1$的假设检验问题. 设$(\widehat{k}_1^0,\ldots,\widehat{k}_l^0)$与$(\widehat{k}_1^*,\ldots,\widehat{k}_{l+1}^*)$分别是原假设与备择假设下变点的估计, 从而定义统计量

$$L(l+1|l) = R(\widehat{k}_1^*,\ldots,\widehat{k}_{l+1}^*) - R(\widehat{k}_1^0,\ldots,\widehat{k}_l^0),$$

其中$\widehat{k}_1^*,\ldots,\widehat{k}_{l+1}^*$中有且仅有$l$个与$\widehat{k}_1^0,\ldots,\widehat{k}_l^0$相同. 定义$R(0)=0$, 则当$l=0$时, 即为至多一个变点的均值变点检测问题. 若原假设$H_0: m=l$成立, 则$\sup L(l+1|l)$是较小的数, 即对一个较小的数c_α, 有$\mathsf{P}\{\sup L(l+1|l)>c_\alpha\}<\alpha$.

定理 3.1　如果假设2.1、2.2以及原假设$H_0: m=l$成立, 则

$$\sup L(l+1|l) \xrightarrow{\mathrm{d}} \max\{\xi_1,\ldots,\xi_{l+1}\}, \tag{3.1}$$

其中$\xi_i=\sup\limits_{\upsilon}\{s_i|\omega_i^0(\upsilon)|+A_{i\upsilon}+B_{i\upsilon}\}$, $s_i=\sqrt{1/(k_i^0-k_{i-1}^0)}$, $\omega_i^0(\upsilon)$为$[0,1]$上的Brown桥, 并且

$$A_{i\upsilon} = \Big\{\frac{(\tau_{i-1}^0-\tau_{i-2}^0)(\upsilon-\tau_{i-1}^0)}{(\upsilon-\tau_{i-2}^0)^2} - \frac{(\tau_{i-1}^0-\tau_{i-2}^0)(\tau_i^0-\tau_{i-1}^0)}{(\tau_i^0-\tau_{i-2}^0)^2}\Big\} f(\zeta_{i,1})\big|\Delta_{[k_{i-2}^0,k_i^0]}\big|,$$

$$B_{i\upsilon} = \Big\{\frac{(\tau_i^0-\upsilon)(\tau_{i+1}^0-\tau_i^0)}{(\tau_{i+1}^0-\upsilon)^2} - \frac{(\tau_i^0-\tau_{i-1}^0)(\tau_{i+1}^0-\tau_i^0)}{(\tau_{i+1}^0-\tau_{i-1}^0)^2}\Big\} f(\zeta_{i,2})\big|\Delta_{[k_{i-1}^0,k_{i+1}^0]}\big|,$$

此处$0<\upsilon<1$, $i=1,\ldots,l+1$, $\zeta_{i,1}$为介于$m_l^*-\mu_{i-1}$与$m_l^*-\mu_i$之间的点, $\zeta_{i,2}$为介于$m_l^*-\mu_i$与$m_l^*-\mu_{i+1}$之间的点, m_l^*为一个只与原假设下变点个数l以及位置有关的常数. 当$i=0$时, $\tau_i^0=0$; 当$i\geq l+1$时, $\tau_i^0=1$.

以下仅对$l=1$的情况由引理予以证明, $l>1$的情况类似可得.

引理 3.1　若假设2.1、2.2以及原假设$H_0: l=1$成立, 则

$$m_T \xrightarrow{\text{a.s.}} m^*,$$

其中m^*为$\tau_1^0 F(\cdot-\mu_1)+(1-\tau_1^0)F(\cdot-\mu_2)$的中位数.

证明:　由Serfling (1992)中的定理6.1可得.　□

引理 3.2　若假设2.1、2.2以及原假设$H_0: l=1$成立, 则对任意的$\varepsilon>0$, 当$T\to\infty$时, 有

$$\mathsf{P}\{|R(k_1^0)-R(\widehat{k}_1^0)|>\varepsilon\}\to 0. \tag{3.2}$$

证明:　类似于Kokoszka和Leipus (2000)可证, 当$\Delta_T\to 0$, $T^{1/2}\Delta_T^2\to\infty$时, $\widehat{\tau}_1^0-\tau_1^0=O_P(T^{-1}\Delta_T^4)$, 此处$\Delta_T=\mu_2-\mu_1$. 从而可以定义集合$K_T(M)=\{k: k=[k_1^0+\upsilon\Delta_T^4], |\upsilon|\le M\}$, 其中$M$为任意正数, 且满足: $M\to\infty$, $M/T\to 0$. 由对称性, 仅考虑$\upsilon>0$的情况. 令$y_t=\mathrm{sgn}(Y_t-m_T)-\mathsf{E}\mathrm{sgn}(Y_t-m_T)$, 则当$k\in K_T(M)$时, 则

$$\begin{aligned}
|R(k)-R(k_1^0)| &\le \Big|\frac{k(T-k)}{T^2}\Big\{\frac{1}{k}\sum_{t=1}^{k}\mathrm{sgn}(Y_t-m_T)-\frac{1}{T-k}\sum_{t=k+1}^{T}\mathrm{sgn}(Y_t-m_T)\Big\}\\
&\quad -\frac{k_1^0(T-k_1^0)}{T^2}\Big\{\frac{1}{k_1^0}\sum_{t=1}^{k_1^0}\mathrm{sgn}(Y_t-m_T)-\frac{1}{T-k_1^0}\sum_{t=k_1^0+1}^{T}\mathrm{sgn}(Y_t-m_T)\Big\}\Big|\\
&\le \frac{1}{T}\Big|\sum_{t=k_1^0}^{k}y_t\Big|+\Big|\frac{k_1^0-k}{T^2}\sum_{t=1}^{T}y_t\Big|\\
&\quad +\frac{k_1^0|k_1^0-k|}{T^2}\Big|\mathsf{E}\mathrm{sgn}(Y_1-m_T)-\mathsf{E}\mathrm{sgn}(Y_{k_1^0+1}-m_T)\Big|\\
&= I_1+I_2+I_3.
\end{aligned}$$

由De Jong和Davidson (2000)的定理3.1知, 当$\Delta_T\to 0$, $T^{1/2}\Delta_T^2\to\infty$时, 有

$$I_1=\frac{1}{T\Delta_T^2}\Big|\Delta_T^2\sum_{t=0}^{\upsilon\Delta_T^{-4}}Y_{t+k_1^0}\Big|=o_P(1).$$

同理可证$I_2=o_P(1)$. 由引理3.1可得

$$\begin{aligned}
I_3 &= \frac{k_1^0|k_1^0-k|}{T^2}|\mathsf{E}\mathrm{sgn}(Y_1-m_T)-\mathsf{E}\mathrm{sgn}(Y_{k_1^0+1}-m_T)|\\
&\to \tau_1^0|\tau_1^0-\tau||F(m^*-\mu_1)-F(m^*-\mu_2)|\\
&= \tau_1^0|\tau_1^0-\tau|f(\zeta)\Delta_T=o(1),
\end{aligned}$$

ζ为介于$m^*-\mu_1$与$m^*-\mu_2$之间的点. 综合可得(3.2).　□

引理 3.3　若假设2.1、2.2以及原假设$H_0: l=1$成立, 则对任意满足$|k_1^0-h|>\eta T$且$h>\eta T$的h, 有

$$R(h,\widehat{k}_1^0)-R(\widehat{k}_1^0)\to s_1|\omega_1^0(\upsilon)|+B_{1\upsilon}; \tag{3.3}$$

和

$$R(\widehat{k}_1^0, h) - R(\widehat{k}_1^0) \to s_2|\omega_2^0(\upsilon)| + A_{1\upsilon}. \tag{3.4}$$

证明:　以下仅证(3.3), (3.4)类似可证. 注意到$R(h, k_1^0) = R_{[0,k_1^0]}(h) + R_{[h,T]}(k_1^0)$,

$$\begin{aligned} & R(h, \widehat{k}_1^0) - R(\widehat{k}_1^0) \\ = \ & R(h, \widehat{k}_1^0) - R(h, k_1^0) + R_{[0,k_1^0]}(h) + (R_{[h,T]}(k_1^0) - R(k_1^0)) + R(k_1^0) - R(\widehat{k}_1^0). \end{aligned}$$

由引理3.2可得, $R(k_1^0) - R(\widehat{k}_1^0) = o_P(1)$, $R(h, \widehat{k}_1^0) - R(h, k_1^0) = o_P(1)$, 结合De Jong和Davidson (2000)的定理3.1, 可得$R_{[0,k_1^0]}(h) \to s_1|\omega_1^0(\upsilon)|$, 又由于$R_{[h,T]}(k_1^0) - R(k_1^0) \to B_{1\upsilon}$, 因此(3.3)成立. 引理证毕.　□

引理 3.4　若假设2.1、2.2以及原假设$H_0: m = 1$成立, 则

$$\sup L(2|1) \to \max\{\xi_1, \xi_2\}. \tag{3.5}$$

证明:　由引理3.2和3.3可得.　□

注记 2　上述$l = 1$时3.1的证明推广到一般情形时, 关键是将引理3.1推广, 而这点可以由Serfling (1992)中的定理2.1, 即关于样本中位数的指数不等式推出.

由Brown桥尾概率公式可得如下结论:

推论 3.1　在定理3.1的条件下, 有

$$\lim_{T\to\infty} \mathsf{P}(\max\{L(l+1|l) - (A_{i\upsilon} + B_{i\upsilon})\}/s_i > c_\alpha) = \prod_{i=1}^{l+1} G_i(c_\alpha), \tag{3.6}$$

其中$G_i(c_\alpha) = 1 - \sum\limits_{k=-\infty}^{\infty} (-1)^k \exp\{-2k^2 c_\alpha^2\}$.

令m_0为实际的变点数, $\widehat{m}$为多变点检验接受原假设时的变点个数, 作为m_0的估计, α_T为样本容量为T时检验的显著水平, 则有以下结论成立:

定理 3.2　在假设2.1、2.2下, 如果$T \to \infty$时, $\alpha_T \to \infty$且$T\alpha_T = O(1)$, 则

$$\mathsf{P}(\widehat{m} = m_0) \to 1. \tag{3.7}$$

证明:　若原假设成立, 即$\widehat{m} = l = m_0$, 则当$T \to \infty$时,

$$\mathsf{P}(\widehat{m} = m_0) = 1 - \alpha_T \to 1.$$

另一方面, 若备择假设成立, 即$\widehat{m} = l + 1 \le m_0$, 则

$$L^* = \max\{L(l+1|l)/s_i - (A_{i\upsilon} + B_{i\upsilon})/s_i\} = O_P(T^{1/2}). \tag{3.8}$$

令c_{α_T}为与α_T对应的临界值, 由推论3.1可得

$$\alpha_T=\prod_{i=1}^{l+1}G_i(c_{\alpha_T})\le 2^{l+1}\exp\Big\{-2\sum_{i=1}^{l+1}c_{\alpha_T^2}\Big\}.$$

从而存在$M>0$, 使得$c_{\alpha_T}\le M(\log(2\alpha_T)^{-1})^{1/2}$, 又由假设$T\alpha_T=O(1)$, 可得$c_{\alpha_T}^2=O(\log(T/2))$, 从而$\mathsf{P}\{L^*>c_{\alpha_T}\}\to 1$. 即以概率1拒绝原假设, 即重新设置原假设进行检验, 直到原假设被接受时, 有$\widehat{m}=l=m_0$, 从而有(3.7)成立. 定理证毕. $\square$

§4. 模拟与实际例子

为了进一步验证本文提出的检验方法的有效性, 我们对以下模拟和实际例子进行分析.

4.1 随机模拟

考虑含有变点的模型

$$Y_t=\mu+\varepsilon_t,\qquad t=1,\ldots,$$

其中新息ε_t分别为独立同分布变量. 用此模型产生样本$Y_t,\ldots,Y_{1500}$, 在样本区间$[1,500]$与$[1001,1500]$上, $\mu=0$, 在$[501,1000]$上, $\mu^*=1$. 考虑含有一个变点的模型(I): Y_t, $t=1,\ldots,1000$与含有两个变点的模型(II): Y_t, $t=1,\ldots,1500$. 为了考察本文统计量对厚尾数据时的统计性质, 新息ε_t分别取标准正态变量、自由度为2的t分布和Cauchy分布.

用5000次试验中拒绝原假设的百分数作为经验势函数值, 模拟结果见表. 对于含有一个变点的模型(I), 在检验水平0.05与0.01都几乎以100%拒绝原假设$H_0: m=0$, 重新设置原假设$H_0: m=1$, 则分别得到检验经验水平值, 如表1、2、3所示. 三种新息, 检验的经验水平值都与真实检验水平相当接近, 其检验的经验势函数值因此接受原假设$H_0: m=1$, 并停止假设检验, 同时认为模型含有一个变点. 同样, 对于模型(II), 接受两个变点的假设. 比较表1与表2、3, 可以看到, 当新息为方差无穷的$t(2)$分布与Cauchy分布, 检验的经验势函数值与新息为正态时的势函数值接近.

为了比较本文的统计量与Bai (1999)中的统计量在新息为厚尾序列时的统计性质, 我们对上述两个模型分别在三种新息下计算得到其经验势函数与经验水平值, 列于表1、2、3中. 表1、2、3括弧中的数据为应用Bai (1999)中的方法对以上两个模型在三种新息下计算得到的经验是函数与经验水平值. 对比两组数据可以看到, 在新息接近于正态分布时, 本文统计量与Bai (1999)中的统计量都具有较高的势与良好的经验水平. 而当新息尾部逐渐变厚时, Bai (1999)中的统计量的经验水平严重偏离检验水平, 而本文统计量的经验水平仍然接近检验水平, 这说明在新息尾部较厚时, 本文所给的方法受到厚尾新息中异常值的影响较小, 也更为有效.

表1　新息为标准正态时的经验水平及经验势函数值

H_0	$m=0$	$m=0$	$m=1$	$m=1$	$m=2$	$m=2$
H_1	$m=1$	$m=1$	$m=2$	$m=2$	$m=3$	$m=3$
检验水平	0.05	0.01	0.05	0.01	0.05	0.01
模型(I)	1(1)	1(1)	0.0490(0.0512)	0.0088(0.0112)		
模型(II)	1(1)	1(1)	1(1)	1(1)	0.0542(0.0494)	0.0136(0.0120)

表2　新息为$t(2)$分布时的经验水平及经验势函数值

H_0	$m=0$	$m=0$	$m=1$	$m=1$	$m=2$	$m=2$
H_1	$m=1$	$m=1$	$m=2$	$m=2$	$m=3$	$m=3$
检验水平	0.05	0.01	0.05	0.01	0.05	0.01
模型(I)	1(1)	1(1)	0.0638(0.8860)	0.0132(0.7590)		
模型(II)	1(1)	1(1)	1(1)	1(1)	0.0666(0.8996)	0.0164(0.7904)

表3　新息为Cauchy分布时的经验水平及经验势函数值

H_0	$m=0$	$m=0$	$m=1$	$m=1$	$m=2$	$m=2$
H_1	$m=1$	$m=1$	$m=2$	$m=2$	$m=3$	$m=3$
检验水平	0.05	0.01	0.05	0.01	0.05	0.01
模型(I)	1(1)	1(1)	0.0770(1)	0.0158(1)		
模型(II)	1(1)	1(1)	1(1)	1(1)	0.0836(1)	0.0234(1)

4.2　应用实例

在金融市场中, 一些突发事件常常会在数据中得到反应. 考虑鲁北化工(LBC)股票(上证600727)从2004年7月1日至2005年12月16日收盘价共357个观测值作实例分析. 数据的原始序列P_t如图1所示. 该股票的对数收益率序列$Y_t=\ln P_t-\ln P_{t-1}$如图2所示. 在对数收益率序列中, 在$k=125$、175、300附近可能存在变点. 设置假设$H_0: l=0$与$H_1: l=1$, 统计量$L(1|0)=0.4861$, 大于显著水平0.05与0.01的对应分位数0.0696和0.0858, 故可以认为至少存在一个变点$\widehat{k}_1=175$. 应用本文中的统计量进一步检验数据在$t=112$以及$t=310$处存在变点, 而在其他地方不存在均值变点. 事实上, $t=112$则对应的实际时间2004年12月10日, 此时公司正在对内部一些不规范之处进行整改, 造成股票价格的下跌. $t=175$对应日期为2005年3月21日, 此前3月9日公司第3届董事会第17次会议发布公告称, 公司2004年度纯利润将比2003年下降50%, 这个负面消息直接导致该公司股票价格连续下跌. $t=310$对应日期为2005年10月12日, 由于在9月29日公司发布2005年第三季度财务拟披露报告, 该报告指出2005年全年纯利润将比2004年降低50%, 该消极因素导致公司股票的第三次大幅变化.

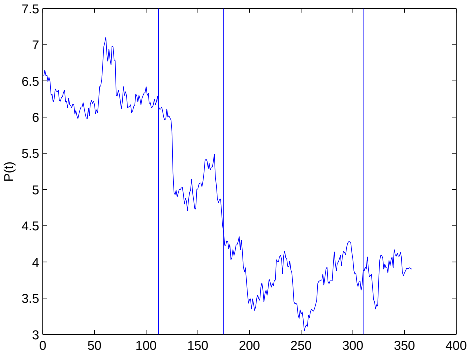

图1 LBC股票价格原始数据图

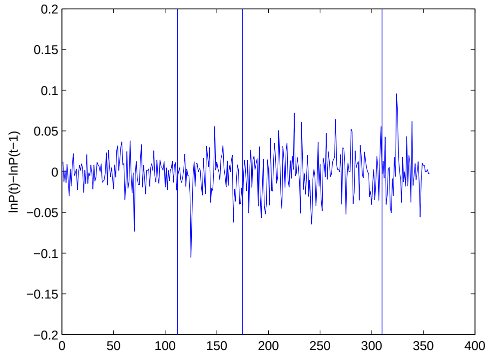

图2 LBC股票对数收益率图

§5. 结 束 语

对独立随机序列的均值多变点检测问题, 本文基于随机序列的中位数构造了一种非参数的检验统计量, 得到统计量的极限分布, 同时得到了变点时刻与变点个数的一致估计. 模拟结果表明, 当新息尾部较厚时, 本文所给的方法具有较好的势与经验水平. 当新息过程存在异常值时, 如何构造均值变点的稳健检测与估计, 本文的方法是否可以推广到其他模型, 是否仍具有良好的统计性质, 将是我们今后研究的重点.

参 考 文 献

[1] Page, E.S., Continuous inspection schemes, *Biometrika*, **41(1-2)**(1954), 100–115.

[2] Csörgő, M. and Horváth, L., *Limit Theorems in Change-Point Analysis*, Chichester: John Wiley & Sons, 1997.

[3] Fang, Y., et al., Online change detection: Monitoring land cover from remotely sensed data, *Sixth IEEE International Conference on Data Mining – Workshops* (*ICDMW'06*), New York: IEEE Press, 2006.

[4] Yu, X., Tan, H. and Wan, W., A novel speaker change detection algorithm, *International Conference on Communications, Circuits and Systems* (*ICCCAS 2007*), New York: IEEE Press, 2007.

[5] Wang, H., Zhang, D. and Shin, K., Change-point monitoring for the detection of DoS attacks, *IEEE Transactions on Dependable and Secure Computing*, **1(4)**(2004), 193–208.

[6] Tsechpenakis, G., et al., Robust online change-point detection in video sequences, *Conference on Computer Vision and Pattern Recognition Workshop* (*CVPRW'06*), New York: IEEE Press, 2006.

[7] 张开鹏, 吴代华, 李卓球, 基于均值变点检验的结构损伤定位方法研究, 第十三届全国结构工程学术会议论文集, 2004.

[8] Bai, J.S., Estimating multiple breaks one at a time, *Econometric Theory*, **13(3)**(1997), 315–352.

[9] Bai, J.S., Likelihood ratio tests for multiple structural changes, *Journal of Econometrics*, **91(2)**(1999), 299–323.

[10] Lavielle, M. and Moulines, E., Least-suqares estimation of an unknown number of shifts in a time series, *Journal of Time Series Analysis*, **21(1)**(2000), 33–59.

[11] 韩四儿, 田铮, 武新乾, 一类股市波动性预测模型的多变点检验, *系统工程理论与实践*, **26(3)**(2006), 94–101.

[12] Meerschaert, M. and Scheffler, H., Nonparametric methods for heavy tailed vector data: A survey with applications from finance and hydrolody, *Recent Advances and Trends in Nonparametric Statistics*, Akritas, M.G. and Politis, D.N. (Editors), Elsevier Science, 2003, 265–279.

[13] Brcich, R., Iskander, R. and Zoubir, A., The stability test for symmetric alpha-stable distributions, *IEEE Transactions on Signal Processing*, **53(3)**(2005), 977–986.

[14] 张晨, 韩东, 宗福季, 检测Lévy稳定过程均值变点的平均运行时间估计, *应用概率统计*, **23(4)**(2007), 384–394.

[15] Shi, X., Wu, Y. and Miao, B., Strong convergence rate of estimators of change point and its application, *Computational Statistics and Data Analysis*, **53(4)**(2009), 990–998.

[16] Serfling, R., Nonparametric confidence intervals for generalized quantile parameters in multi-sample contexts, *Nonparametric Statistics and Related Topics*, Saleh, A.K.Md.E. (Editor), North Holland, 1992, 121–139.

[17] Kokoszka, P. and Leipus, R., Change-point estimation in ARCH models, *Bernoulli*, **6(3)**(2000), 513–539.

[18] De Jong, R.M. and Davidson, J., The functional central limit theorem and weak convergence to stochastic integrals I: weakly dependent and fractionally integrated processes, *Econometric Theory*, **16(5)**(2000), 621–642.

A Nonparametric Test for Multiple Changes in the Mean of Independent Random Series

Qin Ruibing[1,2]　　Tian Zheng[2]　　Chen Zhanshou[2,3]

([1] *School of Mathematical Science, Shanxi University, Taiyuan, 030006*)

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

([3] *Department of Mathematics, Qinghai Normal University, Xining, 810008*)

A nonparametric procedure is proposed to detect multiple changes in the mean of independent random series and the asymptotic distribution is derived. Simultaneously, the estimators for the locations of the change points are obtained. Moreover, the performance of the test is studied by Monte Carlo simulation, which demonstrates that the proposed test has high powers and good sizes for heavy-tailed innovations. Finally, the feasibility of the proposed test is illustrated by the application on the LBC data of stock prices.

Keywords: Independent random series, multiple change points, the nonparametric test, sample median.

AMS Subject Classification: 62D05, 62M10.