Journal of Mathematical Research with Applications
May, 2015, Vol. 35, No. 3, pp. 330–342
DOI:10.3770/j.issn:2095-2651.2015.03.011
Http://jmre.dlut.edu.cn

Monitoring Distributional Changes in Autoregressive Models Based on Weighted Empirical Process of Residuals

Fuxiao LI[1,*], **Zheng TIAN**[1,2], **Zhanshou CHEN**[3]

1. *Department of Applied Mathematics, Northwestern Polytechnical University, Shaanxi* 710129*, P. R. China;*
2. *State Key Laboratory of Remote Sensing Science, Chinese Academy of Science, Beijing* 100101*, P. R. China;*
3. *Department of Mathematics and information, Qinghai Normal University, Qinghai* 810008*, P. R. China*

Abstract Change monitoring of distribution in time series models is an important issue. This paper proposes a procedure for monitoring changes in the error distribution of autoregressive time series, which is based on a weighed empirical process of residuals with weights equal to the regressors. The asymptotic properties of our monitoring statistic are derived under the null hypothesis of no change in distribution. The finite sample properties are investigated by a simulation. As it turns out, the procedure is not only able to detect distributional changes but also changes in the regression coefficient and mean. Finally, we apply the statistic to a groups of financial data.

Keywords distributional changes; autoregressive models; weighted empirical process of residuals

MR(2010) Subject Classification 62F12

1. Introduction

The problem of testing for parameter changes in time series models has attracted much attention from researchers since time series often experience structural changes due to the changes of monetary policy and critical social events. There are two distinctly different approaches to tackle such problems, namely, (1) retrospective (off-line) or a posteriori test, and (2) sequential (on-line) or a priori test. The former tests study a fixed historical sample, for surveys we refer to Csorgo and Horvath [1], Lee et al. [2], Berkes et al. [3], Perron [4] and Qin and Tian [5]. The latter tests on-line monitor new observations to see if a change occurs, for instance, we can mention Chu et al. [6], Leisch et al. [7], Horvath et al. [8] and Chen et al. [9].

While many articles are devoted to detecting parameter change in time series, testing for distributional changes in time series has also received wide attention. Karunamuni and Zhang [10] investigated the detection of a change in distribution for independent observations. Inoue

Received December 10, 2013; Accepted March 9, 2015
Supported by the National Natural Science Foundation of China (Grant No. 11301291) and the Open Fund of State Key Laboratory of Remote Sensing Science of China (Grant No. OFSLRSS201206).
* Corresponding author
E-mail address: lifuxiao1987@sina.com (Fuxiao LI)

[11] suggested nonparametric tests of change in the distribution of a strong mixing sequence. Horvath et al. [12] proposed a test to detect a change in the error distribution of an ARCH sequence. Huskova and Meintanis [13,14] developed the detection problem for distributional changes with independent identically distributed (i.i.d.) observations. Compared with retrospective test, more attention is being paid to sequential detection of a change in the distribution of a time series. Gombay [15] studied sequential detection of distributional changes in i.i.d. observations. Huskova and Chochola [16] investigated sequential detection of a change in distribution in case of independent and dependent observations. Na et al. [17] developed a general monitoring procedure for time series and applied their method to monitor distributional changes in i.i.d. samples. For autoregressive (AR) models, Lee et al. [18] proposed a monitoring procedure related to Kolmogorov-Smirnov (K-S) statistics for an early detection of error distribution changes in AR(p) models. Hlavka et al. [19] discussed a sequential test based on the empirical characteristic function (ECF) of the residuals for monitoring changes in the error distribution. However, the tests proposed by Lee et al. [18] and Hlavaka et al. [19] could not achieve satisfactory effects when the error of AR(p) models changes from normal distribution to heavy-tailed distribution. In other words, their procedures have lower powers and longer detection delays in that case. Heavy-tailed distribution is applied to many fields of probability and statistics, such as branching process, random theory, queuing theory and risk theory and so on（see Su et al. [20]）. Hence, studying for a monitoring scheme, which is sensitive for this type of change, is necessary and meaningful.

With the above considerations in mind, we provide in this paper a monitoring procedure based on a weighted empirical process of residuals with weights equal to the regressors. The weighted empirical process of residuals was employed by Bai [21] in order to test parameter constancy in linear regression models in an off-line setting. Simulations indicate that the proposed procedure has higher powers and shorter detection delays than that of Lee et al. [18] and Halvka et al. [19] when the error terms of AR(p) models change from normal distribution to heavy-tailed distribution, furthermore, this procedure also produces satisfactory results for changes in the regression coefficient and mean.

The rest of the paper is organized as follows. Section 2 introduces the models and assumptions. The test statistic and its asymptotic behaviors under both null and alternative hypothesis are stated in Section 3. In Section 4, we show the finite sample performance through simulations and empirical application. Section 5 concludes the paper. The proof of Theorem 3.1 is gathered in Section 6.

2. Assumptions and models

Let $\{x_t, t = p+1, \ldots\}$ be an AR(p) process defined by the equation

$$x_t = \beta^T X_{t-1} + \varepsilon_t \tag{1}$$

where $X_{t-1} = (x_{t-1}, \ldots, x_{t-p})^T$, and $\beta = (\beta_1, \ldots, \beta_p)^T$ is an unknown regression parameter. The errors ε_t are independent, each having a corresponding distribution function F_t with mean zero and finite variance.

In this paper, we monitor the distributional change in autoregressive models with at most one change point. For some known $T < \infty$, we are interested in testing the hypothesis

$$H_0 : F_t = F_0,\ t = 1, \dots, T, \dots, \lfloor T\kappa \rfloor, \tag{2}$$

$$H_1 : F_t = F_0,\ T < t < T + k^*;\ F_t = F^0,\ T + k^* \le t \le \lfloor T\kappa \rfloor, \tag{3}$$

where the distribution functions $F_0 \neq F^0$ as well as the time of change k^* are assumed unknown, κ is some fixed number larger than 1, $\lfloor\cdot\rfloor$ denotes the integer part.

In the hypothesis-testing problem exemplified by H_i, $i = 0, 1$, suppose that there exists a fixed set $x_1, \dots, x_T$ of historic data which involve no change, i.e., $F_1 = \dots = F_T$. Based on this data set, we complete the estimator $\hat{\beta}_T := \hat{\beta}(x_1, \dots, x_T)$ of β in model (1). Then, in view of the fact that the errors are unobserved, typically one calculates the residuals

$$\hat{\varepsilon}_t = x_t - \hat{\beta}'_T X_{t-1},$$

and it is on the basis of these residuals that the null hypothesis H_0 will be tested against H_1.

Now suppose that we are operating with an on-line monitoring scheme. So that the test statistic, say S_t, is computed sequentially at each time point, and that the null hypothesis should be rejected when the value of the statistic exceeds an appropriately chosen constant c_α for the first time. Otherwise we should continue monitoring. The associated stopping rule is given by

$$\tau(T) = \begin{cases} \inf\{T < t < \lfloor T\kappa \rfloor : S_t > c_\alpha\}, \\ \lfloor T\kappa \rfloor,\ \text{if } S_t \le c_\alpha \text{ for } T < t < \lfloor T\kappa \rfloor. \end{cases}$$

As in classical hypothesis testing, our aim is to control the overall value of α,

$$\lim_{T\to\infty} P_{H_0}(\tau(T) < \infty) = \alpha, \tag{4}$$

$$\lim_{T\to\infty} P_{H_1}(\tau(T) < \infty) = 1, \tag{5}$$

the probability $\alpha \in (0, 1)$ controls the false alarm rate, (4) ensures that the probability of false alarm is asymptotically bounded by α, while (5) means that a change is detected with probability approaching one.

In the remainder of this section, we state the assumptions on the regression model

Assumption 2.1 *$\{\varepsilon_t, t = 1, 2, \dots\}$ are i.i.d. random variables with common distribution F_0 having zero mean, positive variance and $E|\varepsilon_t^4| < \infty$, F_0 admits a density function f, $f > 0$. Both $f(x)$ and $xf(x)$ are assumed to be uniformly continuous on the real line. Furthermore, there exists a finite number L such that $|xf(x)| < L$ and $|f(x)| < L$ for all x.*

Assumption 2.2 *The initial values $x_1, x_2, \dots, x_p$ are independent of $\varepsilon_{p+1}, \dots, \varepsilon_T$, let $\beta_p \neq 0$, and the roots of the polynomial $z^p - \beta_1 z^{p-1} - \dots - \beta_p$ are less than one in absolute value.*

Assumption 2.3 *$(Y^T Y)^{1/2}(\hat{\beta}_T - \beta) = O_p(1)$, as $T \to \infty$, where $Y = (x_1, x_2, \dots, x_T)^T$.*

Assumption 2.4 *The regressors satisfy*

$$\lim_{T\to\infty} \max_{1\le t\le \lfloor T\kappa\rfloor} \frac{1}{T^{1/2}}|x_t| = o_p(1),$$

$$\lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{\lfloor Ts\rfloor} x_t^2 = \lim_{T\to\infty} \frac{1}{T}E\sum_{t=1}^{\lfloor Ts\rfloor} x_t^2 = l(s), \text{ uniformly s} \in [1,\kappa],$$

where $l(s)$ is positive for $s > 0$.

Assumption 2.5 *For every fixed s_1, there exists a sequence of positive number $z_T = O_p(1)$ such that*

$$\frac{1}{T}\sum_{t=\lfloor Ts_1\rfloor}^{\lfloor Ts\rfloor} |x_t| \le (s-s_1)z_T \quad \text{a.s.},$$

for all $s \ge s_1$, and the tail probability of z_T satisfies, for some $\rho > 0$,

$$P(z_T > C) \le M/C^{2(1+\rho)}$$

where $C > 0$ and $M > 0$.

Assumption 2.6 *There exist $\gamma > 1$, $\alpha > 1$ and $K < \infty$ such that for all $0 \le s' \le s'' \le 1$, and for all T,*

$$\frac{1}{T}\sum_{i<t<j} E(x_t^2)^\gamma \le K(s''-s') \text{ and } E\Big(\frac{1}{T}\sum_{i<t<j} x_t^2\Big)^\gamma \le K(s''-s')^\alpha,$$

where $i = \lfloor Ts'\rfloor$, $j = \lfloor Ts''\rfloor$.

Assumption 2.7 *For arbitrary T, there exist $\delta > 0$ and $M < \infty$ such that*

$$E\Big(\frac{1}{T}\sum_{t=1}^{T}|x_t|^{3(1+\delta)}\Big) < M \text{ and } E\Big(\frac{1}{T}\sum_{t=1}^{T}|x_t|^3\Big)^{1+\delta} < M.$$

Remark 2.8 Assumptions 2.1 and 2.2 are sufficient conditions for stationary AR(p) process. When the error terms are i.i.d. and have finite variance, then the least squares estimator $\hat\beta_T$ satisfies assumption 2.3. Assumptions 2.4–2.7 are required in the proof of limiting distribution to follow.

3. Main results

Let $\hat\beta_T$ be the least squares estimator β based on the observations $x_1,\dots,x_T$, namely,

$$\hat\beta_T = \Big(\sum_{t=p+1}^{T} X_{t-1}X_{t-1}'\Big)^{-1}\sum_{t=p+1}^{T} X_{t-1}x_t,$$

and calculate the residuals

$$\hat\varepsilon_t = x_t - \hat\beta_T' X_{t-1}, \quad t \ge p+1. \tag{6}$$

For $k \ge p+1$, set

$$\hat F_k(x) = \frac{1}{k}\sum_{i=p+1}^{k} I(\hat\varepsilon_i \le x), \quad -\infty < x < \infty.$$

Let $Y_k = (x_1, \ldots, x_k)^T$ and $A_k = (Y^TY)^{-1/2}(Y_k^TY_k)(Y^TY)^{-1/2}$. Define Γ_T

$$\Gamma_T(\frac{k}{T}, x) = (Y^TY)^{-1/2}\sum_{i=1}^{k} x_i\{I(\hat{\varepsilon}_i \le x) - \hat{F}_T(x)\} - A_k(Y^TY)^{-1/2}\sum_{i=1}^{T} x_i\{I(\hat{\varepsilon}_i \le x) - \hat{F}_T(x)\}$$

and the test statistic

$$M_T = \max_{T<k\le\lfloor T\kappa\rfloor} \sup_{-\infty<x<+\infty} |\Gamma_T(\frac{k}{T}, x)|. \tag{7}$$

The statistic Γ_T, obtained by constructing a weighted empirical process of residuals with weights equal to the regressors, was employed by Bai [21] to detect parameter change in linear regression models.

Let $B(u, v)$ be a Gaussian process on $[1, \kappa] \times [0, 1]$ with zero mean and covariance function

$$E\{B(s, u)B(t, v)\} = (\min(s, t) - st)(\min(u, v) - uv),$$

which we shall call a two parameter Brownian bridge on $[1, \kappa] \times [0, 1]$.

Theorem 3.1 *Under model (1) and Assumptions 2.1–2.7,*

$$\Gamma_T(\frac{\lfloor Ts\rfloor}{T}, \cdot) \Rightarrow B(\cdot, F(\cdot)),$$

where $B(\cdot, F(\cdot))$ *is a two parameter Brownian bridge on* $[1, \kappa] \times [0, 1]$, $\Rightarrow$ *denotes the weak convergence in the space of* $[1, \kappa] \times [0, 1]$.

Corollary 3.2 *Under model(1) and Assumptions 2.1–2.7,*

$$\lim_{T\to\infty} P(M_T \le x) = G(x), \tag{8}$$

where $G(x)$ *denotes the distribution of the random variable* $\sup_{1<u\le\kappa}\sup_{0\le v\le 1}|B(u, v)|$.

The proof of the theorem is based on the limiting behavior of the process K_T,

$$K_T(s, x) = (X'X)^{-1/2}\sum_{i=1}^{\lfloor Ts\rfloor} X_i\{I(\hat{\varepsilon}_i \le x) - F(x)\},$$

which we call the weighted sequential empirical process of residuals, then

$$\Gamma_T(\frac{\lfloor Ts\rfloor}{T}, x) = K_T(s, x) - A_{\lfloor Ts\rfloor}K_T(1, x).$$

4. Simulations and an application

4.1. Simulations

In this section we report the results of a simulation study that is performed in order to check the finite sample performance of the monitoring procedure considered in the previous section.

The critical values c were obtained by applying a classical bootstrap based on the residuals of the training sample to approximate the limit distribution as indicated in (8). The classical bootstrap based on the estimated residuals of the training sample is as follows. Let $U_T(p + 1), \ldots, U_T(L_T)$ be i.i.d. uniform on $p+1, \ldots, T$ independent of $\{x_t\}$, $\varepsilon^*(t) = \hat{\varepsilon}_{U_T}(t)$ with $\hat{\varepsilon}_j$ as in

(6), where $L_T = \lfloor T\kappa \rfloor$. The bootstrap critical value $c_\alpha(x_1, \ldots, x_T)$ is chosen minimal such that

$$P_T^*(M_T(\varepsilon^*(1), \ldots, \varepsilon^*(T)) \leq c_\alpha(x_1, \ldots, x_T)) \geq 1 - \alpha,$$

where $P_T^*(\cdot) = P(\cdot|x_1, \ldots, x_T)$. We can easily simulate the above conditional distribution by drawing B random realizations of $\{U_T(\cdot)\}$.

The training sample is an AR(1) process with the regression coefficient $\beta = 0.1, 0.5$ and standard normally distributed error terms. The length of training sample $T = 500$ and the monitoring length $(\kappa - 1)T$ with $\kappa = 3$. All experiments were repeated 2000 times, and the bootstrap approximation uses $B = 2000$ replicates. All results are obtained for the level $\alpha = 0.05$. In each simulation, 100 initial observations are discarded to remove initialization effects. Consider the following alternative hypothesis:

(1) Change in mean;

(2) Change in regression coefficient;

(3) Change from $N(0,1)$ to $N(0,2)$;

(4) Change from $N(0,1)$ to χ^2 distribution with 4 degrees of freedom $\chi^2(4)$;

(5) Change from $N(0,1)$ to Cauchy distribution $t(1)$;

(6) Change from $N(0,1)$ to Student t-distribution with 4 degrees of freedom $t(4)$;

(7) Change from $N(0,1)$ to Log-normal distribution Log-$N(\mu, \sigma^2)$ with $\mu = 0$ and $\sigma = 1$.

Next, we evaluate the performance of the K-S, ECF and M_T statistics under the null and alternative hypothesis. In particular, the K-S and ECF statistics were proposed by Lee et al. [18] and Hlavka et al. [19], respectively, M_T is defined by (7). For the ECF statistic we set $\gamma = 1$, $\omega(u) = \exp(-au^2)$ and $a = \hat{\sigma}_T^2/2$, where $\hat{\sigma}_T$ is the standard deviation obtained from the training sample.

Table 1 presents the empirical sizes for K-S, ECF and M_T statistics when the regression coefficient β is 0.1 and 0.5. From the table it can be seen that the empirical sizes are near the nominal size. The same is also true for other βs such as 0.9 although not reported here for brevity.

β	$K-S$	ECF	M_T
0.1	0.065	0.049	0.01
0.5	0.068	0.045	0.03

Table 1 Empirical sizes when $T = 500$, $\kappa = 3$.

Tables 2–8 summarize the empirical powers and some elementary statistics concerning the detection delays when the change occurs at $T + k^*$ with $k^* = 50, 200$. We only report the result for the case $\beta = 0.5$ since the results for other βs are similar to this case. First consider $k^* = 50$, the K-S statistic has the shorter detection delays than the ECF and M_T statistics when the mean shifts, the ECF statistic behaves best when the error terms change from $N(0,1)$ to $N(0,2)$, as shown in Tables 2 and 4. Table 3 and Tables 5–8 show results concerning change in the regression coefficient and change in the distribution of the errors: from $N(0,1)$ to skew distribution ($\chi^2(4)$, Log-$N(0,1)$) and from $N(0,1)$ to heavy-tailed distribution ($t(1)$, $t(4)$), M_T performs best than K-S and ECF statistics among those alternative hypothesis. For instance, the mean detection delays of the K-S statistic, ECF statistic and M_T are, respectively, 873.7, 693.3 and 267.14 when the error distribution changes from $N(0,1)$ to $t(4)$.

k^*		Power	Mean	S.E.	Min	Q1	Median	Q3	Max
	K-S	1.00	29.96	33.50	1	18	26	37	950
50	ECF	1.00	74.37	18.81	30	61	72	85	163
	M_T	1.00	75.80	21.08	20	61	74	89	156
	K-S	0.67	349.64	347.00	6	67	120	800	800
200	ECF	1.00	95.58	35.19	10	71	93	117	288
	M_T	1.00	73.22	23.28	13	57	71	88	170

Table 2 Empirical powers and summary of the detection delays when a change in the mean of +1 occurs, $T = 500$, $\kappa = 3$, $k^* = 50, 200$.

k^*		Power	Mean	S.E.	Min	Q1	Median	Q3	Max
	K-S	0.46	532.69	446.77	1	48	950	950	950
50	ECF	0.73	600.45	283	35	359	580.5	950	950
	M_T	1.00	220.05	91.57	23	160.5	215.5	292	581
	K-S	0.08	778.48	124.11	11	800	800	800	800
200	ECF	0.44	656.10	218.58	13	528	800	800	800
	M_T	1.00	203.57	86.74	10	141	193.5	255	518

Table 3 Empirical powers and summary of the detection delays when the regression coefficient β change from 0.5 to 0.8, $T = 500$, $\kappa = 3$, $k^* = 50, 200$.

k^*		Power	Mean	S.E.	Min	Q1	Median	Q3	Max
	K-S	0.27	729.14	377.32	1	227.5	950	950	950
50	ECF	1.00	202.87	68.61	54	156	192	238	567
	M_T	1.00	327.62	115.72	123	256	306	385	683
	K-S	0.06	797.99	39.03	3	800	800	800	800
200	ECF	1.00	246.69	99.56	13	177	236	302	800
	M_T	1.00	358.27	137.19	53	259	343	447	800

Table 4 Empirical powers and summary of the detection delays when the errors distribution changes from $N(0,1)$ to $N(0,2)$, $T = 500$, $\kappa = 3$, $k^* = 50, 200$.

k^*		Power	Mean	S.E.	Min	Q1	Median	Q3	Max
	K-S	1.00	7.80	4.24	1	5	7	10	38
50	ECF	1.00	60.29	11.65	23	52	59	68	109
	M_T	1.00	6.28	2.50	1	4	6	8	19
	K-S	1.00	40.03	42.43	2	16	28	50	800
200	ECF	1.00	76.58	20.83	12	62	76	90	141
	M_T	1.00	6.31	2.58	1	4	6	8	18

Table 5 Empirical powers and summary of the detection delays when the error distribution change from $N(0,1)$ to $\chi^2(4)$, $T = 500$, $\kappa = 3$, $k^* = 50, 200$.

k^*		Power	Mean	S.E.	Min	Q1	Median	Q3	Max
	K-S	1.00	31.53	87.67	1	13	19	28.75	950
50	ECF	1.00	133.24	90.29	28	82	110	152	950
	M_T	1.00	25.85	12.60	1	17	25	34	73
	K-S	0.42	526.28	360.25	4	61	800	800	800
200	ECF	0.97	205.1	170.29	5	95	149	244	800
	M_T	1.00	25.59	13.03	1	16	24	34	78

Table 6 Empirical powers and summary of the detection delays when the error distribution change from $N(0,1)$ to Log-$N(0,1)$, $T = 500$, $\kappa = 3$, $k^* = 50, 200$.

k^*		Power	Mean	S.E.	Min	Q1	Median	Q3	Max
	K-S	0.90	162.22	289.18	1	16	46	120	950
50	ECF	1.00	129.01	35.44	40	103	126	150	282
	M_T	1.00	17.46	13.99	1	6	14	25	98
	K-S	0.37	563.3	346.8	1	89	800	800	800
200	ECF	1.00	158.62	48.81	18	125	155	190	398
	M_T	1.00	17.73	14.10	1	7	14	25	89

Table 7 Empirical powers and summary of the detection delays when the error distribution change from $N(0,1)$ to $t(1)$, $T = 500$, $\kappa = 3$, $k^* = 50, 200$.

k^*		Power	Mean	S.E.	Min	Q1	Median	Q3	Max
	K-S	0.10	873.7	250.62	1	950	950	950	950
50	ECF	0.65	693.3	247.53	78	479.5	728	950	950
	M_T	1.00	267.14	120.31	8	169	254.5	348.5	646
	K-S	0.06	797.2	45.71	9	800	800	800	800
200	ECF	0.42	669.15	187.55	44	554	800	800	800
	M_T	1.00	302.67	151.61	4	194	281	394.5	800

Table 8 Empirical powers and summary of the detection delays when the error distribution change from $N(0,1)$ to $t(4)$, $T = 500$, $\kappa = 3$, $k^* = 50, 200$.

In contrast with $k^* = 50$, the mean detection delays of K-S and ECF statistics increase when the change occurs at $k^* = 200$, for example, prolongs by around 500 and 70 for change in distribution of error from $N(0,1)$ to Log-$N(0,1)$, respectively, as shown in Table 6. However, the location of the change has no significant effect on M_T, even for $k^* = 200$, the change of elementary statistics, such as mean and median, is small.

In conclusion, if there exists changes in mean, the K-S statistic is recommended, and if the error terms change from $N(0,1)$ to $N(0,2)$, the ECF statistic is best. For change in the regression coefficient and changes in the distribution of the errors: From $N(0,1)$ to skew distribution ($\chi^2(4)$, Log-$N(0,1)$) as well as to heavy-tailed distribution ($t(1)$, $t(4)$), M_T has the shorter detection delays than K-S and ECF statistics. Overall, the proposed monitoring statistic M_T performs best.

4.2. Empirical application

In this section, we illustrate our procedures by a group of financial series, which contains 600 Dow Jones Indexes data from March 7 in 2006 to July 23 in 2008 (see Figure 1). The data are processed with first order difference, centered and standardized first (see Figure 2), then using the weighted Kolmogorov-Smirnov test of Bai in [21], there exists a change in the processed data at point 320. The first 320 data follows the model

$$x_t = 0.027x_{t-1} + \varepsilon_1, \quad \varepsilon_1 \sim N(0, 0.379),$$

the last 280 data follows the model

$$x_t = -0.132x_{t-1} + \varepsilon_2, \quad \varepsilon_2 \sim N(0, 1.49),$$

where x_t denotes the processed data. Let initial 200 processed data be the training samples and we monitor from the 201st data at $\alpha = 0.05$ level using K-S, ECF and M_T statistics. As a result, the detection delays of K-S, ECF and M_T statistics are 400 (not detected), 46 and 32. This result indicates that M_T statistic behaves better than K-S and ECF statistics, which is in accordance with our conclusion.

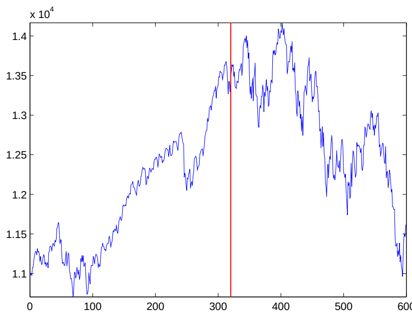

Figure 1 Dow Jones Indexes data

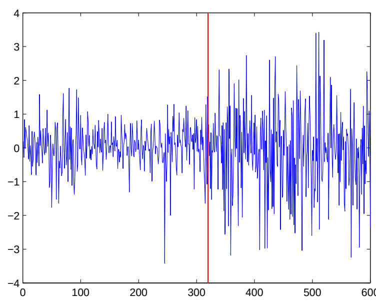

Figure 2 First order difference data

5. Conclusions

In this paper, a monitoring procedure for distributional changes of errors in AR(p) models is proposed. The test is based on a weighed empirical process of residuals with weights equal to the regressors. The asymptotic distribution of the monitoring statistic is derived. A classical bootstrap procedure is applied to approximate the null distribution of the test statistic. The simulation study included K-S and ECF statistics, the results of simulation study suggest that:

(1) K-S statistic has the shorter detection delays than M_T and ECF statistics against "change in mean".

(2) ECF statistic has the shorter detection delays than M_T and K-S statistics against "change from $N(0,1)$ to $N(0,2)$".

(3) M_T performs better than K-S and ECF statistics for "c+hange in regression coefficient, change from $N(0,1)$ to $\chi^2(4)$, change from $N(0,1)$ to Log-normal distribution Log-$N(0,1)$, change from $N(0,1)$ to $t(1)$ and change from $N(0,1)$ to $t(4)$".

It seems that the proposed statistic is recommended in most situations.

6. Mathematical proofs

In this section we state six lemmas needed in proving Theorem 3.1.

Lemma 6.1 *Let $U_1, U_2, \ldots, U_n$ be a sequence of i.i.d. uniformly distributed random variables on $[0,1]$ and x_i $(i = 1, 2, \ldots, T)$ be a sequence of random vectors satisfying assumptions 2.5 and 2.6. Assume that U_i is independent of x_j for $j \le i$. Then the process $Z_T(s,u)$ defined as*

$$Z_T(s,u) = T^{-1/2} \sum_{t=1}^{\lfloor Ts \rfloor} x_t \{I(U_t \le u) - u\}$$

with $Z_T(0,u) = Z_T(s,0) = 0$ is tight in $[1,\kappa] \times [0,1]$.

The process Z_T is a multiparameter process. Lemma 6.1 holds for arbitrary i.i.d. random variables ε_t, in this case $I(U_t \le u) - u$ is replaced by $I(\varepsilon_t \le x) - F(x)$. The proof of Lemma 6.1 is similar to that of Theorem A.1 in Bai [21].

Lemma 6.2 *Under Assumptions 2.1, 2.3, 2.5 and 2.6, the process H_T defined as*

$$H_T(s,x) = (Y^T Y)^{-/2} \sum_{t=1}^{\lfloor Ts \rfloor} x_t \{I(\varepsilon_t \le x) - F(x)\},$$

converges weakly to a Gaussian process with zero mean and covariance function

$$E\{H(r,x)H(s,y)'\} = l(r \wedge s)/l(1)[F(x \wedge y) - F(x)F(y)]. \tag{9}$$

Proof Let $U_i = F(\varepsilon_i)$. Then

$$H_T(s,x) = (Y'Y/T)^{-1/2} Z_T(s, F(x)).$$

Since $(Y'Y/T)$ converges in probability to $l(1)$, the tightness of H_T follows from Lemma 6.1. It is easy to see that $H_T(s,x)$ converges to a normal distribution. To verify the covariance function, consider for $r < s$ and $u = F(x) < v = F(y)$, then by the martingale property, we have that

$$E\{Z_T(r,u)Z_T(s,v)\} = \frac{1}{T}E\Big(\sum_{t=1}^{\lfloor Tr \rfloor} x_t^2\Big)(u - uv),$$

which tends to $l(r)(u - uv)$. The proof is completed. □

Lemma 6.3 *Under Assumptions 2.1, 2.3, 2.5 and 2.6, the process* V_T *defined as*

$$V_T(s,x) = H_T(s,x) - A_{\lfloor Ts \rfloor} H_T(1,x),$$

converges weakly to a Gaussian process V *with mean zero and covariance matrix*

$$E\{V(r,u)V(s,v)\} = \{A(r \wedge s) - A(r)A(s)\}\{u \wedge v - uv\}. \tag{10}$$

Proof The tightness of V_T follows form the tightness of H_T and the convergence of $A_{\lfloor Ts \rfloor}$ to a deterministic function $A(s)$ uniformly in s, by Lemma 6.2,

$$V(s,x) = H(s,x) - A(s)H(1,x),$$

and (10) follows easily from (9). The proof is completed. □

Next, we study the asymptotic behavior of the residual empirical process. Under model (1), $\hat{\varepsilon}_t \le z$ if and only if $\varepsilon_t \le z + (\hat{\beta}_T - \beta)' X_{t-1}$, thus K_T is given by

$$K_T(s,z) = (Y^T Y)^{-1/2} \sum_{t=1}^{\lfloor Ts \rfloor} x_t\{I(\varepsilon_t \le z + (\hat{\beta}_T - \beta)^T X_{t-1}) - F(z)\}.$$

Let $a = (a_1, a_2, \ldots, a_T)$, $b = (b_1, \ldots, b_T)$ be two $1 \times T$ random vectors, and $C = (c_1, c_2, \ldots, c_T)^T$ be a $T \times q$ random matrix ($q \ge 1$). Define

$$Q_T(s,z,a,b) = (C'C)^{-/2} \sum_{t=1}^{\lfloor Ts \rfloor} c_t\{I(\varepsilon_t \le z(1 + a_t T^{-1/2}) + b_t T^{-1/2}) - F(z)\}.$$

For $c_t = x_t$, $a_t = 0$, and $b_t = T^{1/2}(\hat{\beta}_T - \beta)' X_{t-1}$, we have

$$Q_T(s,z,a,b) = K_T(s,z),$$

and moreover,

$$Q_T(s,z,0,0) = H_T(s,z).$$

Assume

(i) The variable ε_t is independent of $\mathcal{F}_{t-1}$, where

$$\mathcal{F}_{t-1} = \sigma - \text{field}\{a_{s+1}, b_{s+1}, c_{s+1}, \varepsilon_s; s \le t-1\}.$$

(ii) $T^{-1}\sum_{t=1}^{T} \|c_t\| = O_p(1)$.

(iii) $T^{-1/2}\max_{1 \le i \le T} |\eta_i| = o_p(1)$, for $\eta_i = a_i, b_i$.

(iv) There exist a $\gamma > 1$ and $A < \infty$ such that for all T

$$E\Big\{\frac{1}{T}\sum_{t=1}^{T} |c_t|^2(|a_t| + |b_t|)\Big\}^\gamma < A \text{ and } \frac{1}{T}\sum_{t=1}^{T} E\{|c_t|^2(|a_t| + |b_t|)\}^\gamma < A.$$

(v) Conditions (iii) and (iv) with $|b_t|$ replaced by $|x_t|$. □

Lemma 6.4 *Under Assumptions 2.1 and (i)–(v),*

$$K_T(s,z,a,b) = K_T(s,z,0,0) + (C^TC/T)^{1/2} f(z) z\Big(\frac{1}{T}\sum_{t=1}^{\lfloor Ts\rfloor} c_t a_t\Big) + f(z)\Big(\frac{1}{T}\sum_{t=1}^{\lfloor Ts\rfloor} c_t b_t) + o_p(1),$$

where the $o_p(1)$ is uniform in s and in z, and for $b_t = x_t\alpha$, the $o_p(1)$ is also uniform in $\alpha \in D$, an arbitrary compact set of R. In particular, the result holds for $b_t = T^{1/2}(\hat\beta_T - \beta)^T X_{t-1}$ as long as $T^{1/2}(\hat\beta_T - \beta) = O_p(1)$.

Lemma 6.5 *Define*

$$Z_T(s,z,a,b) = \frac{1}{\sqrt{T}}\sum_{t=1}^{\lfloor Ts\rfloor} c_t\{I(\varepsilon_t \le z(1+a_tT^{-1/2}) + b_tT^{-1/2}) - F(z(1+a_tT^{-1/2}) + b_tT^{-1/2})\}.$$

(a) Under Assumptions 2.1 and (i)–(iv),

$$\sup_{1<s<\kappa, z\in R} \|Z_T(s,z,a,b) - Z_T(s,z,0,0)\| = o_p(1).$$

(b) Let $b_t = x_t\alpha$ in a compact set D of R for $p \ge 1$ and denote $b(\alpha) = (x_1\alpha, \dots, x_T\alpha)$. Then under Assumption 2.1, (i), (ii) and (v),

$$\sup_{\alpha\in D}\sup_{1\le s\le\kappa, z\in R} \|Z_T(s,z,a,b(\alpha)) - Z_T(s,z,0,0)\| = o_p(1).$$

(c) Let $a_t = r_t'\xi$, $r_t \in R^l$ for some $l \ge 1$; $\xi \in S$, a compact set. Denote $a(\xi) = (r_1'\xi, \dots, r_T'\xi)$, assume (3) and (4) hold with $|a_t| = \|r_t\|$, then under Assumption 2.1, (i) and (ii),

$$\sup_{\xi\in S}\sup_{\alpha\in D}\sup_{1\le s\le\kappa, z\in R} \|Z_T(s,z,a(\xi),b(\alpha)) - Z_T(s,z,0,0)\| = o_p(1).$$

Lemma 6.6 *Under Assumption 2.1 and (i)–(iv), for every $d \in (0, 1/2)$,*

$$\sup_{y,z}\frac{1}{\sqrt{T}}\sum_{t=1}^{T} c_t\|F(y_t^*) - F(z_t^*)\| = o_p(1),$$

where $y_t^* = y(1+a_tT^{-1/2}) + b_tT^{-1/2}$, $z_t^* = z(1+a_tT^{-1/2}) + b_tT^{-1/2}$ and the supreme extends over all pairs of (y,z) such that $|F(y) - F(z)| \le T^{-1/2} - d$.

The proofs of Lemmas 6.4, 6.5 and 6.6 are similar to those of Theorems A.2, A.3 and Lemma A.3 in Bai [21].

Proof of Theorem 3.1 Since $\hat\varepsilon_t = \varepsilon_t - X_{t-1}'(\hat\beta_T - \beta)$, $\hat\varepsilon_t \le z$ if and only if $\varepsilon_t \le z + X_{t-1}'(\hat\beta_T - \beta)$. Applying Lemma 6.4 with $c_t = x_t$, $a_t = 0$ and $b_t = T^{1/2}X_{t-1}'(\hat\beta_T - \beta)$, by

$$Q_T(s,z,a,b) = K_T(s,z) \text{ and } Q_T(s,z,0,0) = H_T(s,z),$$

we have that

$$K_T(s,z) - A_{\lfloor Ts\rfloor}K_T(1,z) = H_T(s,z) - A_{\lfloor Ts\rfloor}H_T(1,z) + \tag{11}$$

$$f(z)(Y'Y/T)^{-1/2}\frac{1}{T}\sum_{t=1}^{\lfloor Ts\rfloor} x_t b_t - f(z)A_{\lfloor Ts\rfloor}(Y'Y/T)^{-1/2}\frac{1}{T}\sum_{t=1}^{T} x_t b_t + \tag{12}$$

$$o_p(1), \tag{13}$$

for $b_t = T^{1/2}X'_{t-1}(\hat{\beta}_T - \beta)$, (12) is identically zero for all $s \in [1, \kappa]$, so the drift terms $K_T(s, z)$ are cancelled out and the limiting distribution of $\Gamma_T(\frac{\lfloor Ts \rfloor}{T}, \cdot)$ is obtained by Lemma 6.3. The proof is completed. □

Acknowledgements We are grateful to the editors and referees for useful suggestions and helpful comments for improving the paper.

References

[1] M. CSORGO, L. HORVATH. *Limit Theorems in Change-Point Analysis*. Wiley, New York, 1997.

[2] S. LEE, J. HA, O. NA, et al. *The CUSUM test for parameter change in time series models*. Scand. J. Statist., 2003, **30**(4): 781–796.

[3] I. BERKES, G. EDIT, L. HORVATH, et al. *Sequential change-point detection in GARCH(p, q) models*. Econometric Theory, 2004, **20**(6): 1140–1167.

[4] P. PERRON. *Dealing with Structural Breaks*. Palgrave Handbook of Econometrics, 2006, **1**: 278–352.

[5] Ruibing QIN, Zheng TIAN. *Bootstrap test for stationarity of heavy-tailed series with structural breaks*. J. Math. Res. Exposition, 2010, **30**(6): 1015–1022.

[6] C. J. CHU, M. STINCHCOMBE, H. WHITE. *Monitoring structural change*. Econometrica, 1996, **64**(5): 1045–1065.

[7] F. LEISCH, K. HORNIK, C. M. KUAN. *Monitoring structural changes with the generalized fluctuation test*. Econometric Theory, 2000, **16**(06): 835–854.

[8] L. HORVATH, M. HUSKOVA, M. KOKOSZKA, et al. *Monitoring changes in linear models*. J. Statist. Plann. Inference, 2004, **126**(1): 25–51.

[9] Zhanshou CHEN, Zheng TIAN, Ruibing QIN, et al. *Sequential monitoring variance change in linear regression model*. J. Math. Res. Exposition, 2010, **30**(4): 610–618.

[10] J. KARUNAMUNI, Shunpu ZHANG. *Empirical Bayes detection of a change in distribution*. Ann. Inst. Statist. Math., 1996, **48**(2): 229–246.

[11] A. INOUE. *Testing for distributional change in time series*. Econometric Theory, 2001, **17**(1): 156–187.

[12] L. HORVATH, P. KOKOSZKA, G. TEYSSIERE. *Empirical process of the squared residuals of an ARCH sequence*. Ann. Statist., 2001, **29**(2): 445–469.

[13] M. HORVATH, S. G. MEINTANIS. *Change point analysis based on empirical characteristic functions*. Metrika, 2006, **63**(2): 145–168.

[14] M. HORVATH, S. G. MEINTANIS. *Change point analysis based on empirical characteristic functions of ranks*. Sequential Anal., 2006, **25**(4): 421–436.

[15] E. GOMBAY. *U-Statistics in sequential tests and change detection*. Sequential Anal., 2004, **23**(2): 257–274.

[16] M. HUSKOVA, O. CHOCHOLA. *Simple Sequential Procedures for Change in Distribution*. Inst. Math. Statist., Beachwood, OH, 2010.

[17] O. NA, Y. LEE, S. LEE. *Monitoring parameter change in time series models*. Stat. Methods Appl., 2011, **20**(2): 171–199.

[18] S. LEE, Y. LEE, O. NA. *Monitoring distributional changes in autoregressive models*. Comm. Statist. Theory Methods, 2009, **38**(16-17): 2969–2982.

[19] Z. HLAVKA, M. HUSKOVA, C. KIRCH, et al. *Monitoring changes in the error distribution of autoregressive models based on Fourier methods*. TEST, 2012, **21**(4): 605–634.

[20] Chun SU, Zhishui HU, Yu CHEN, et al. *A wide class of heavy-tailed distributions and its applications*. Front. Math. China, 2007, **2**(2): 257–286.

[21] Jushan BAI. *Testing for parameter constancy in linear regressions: empirical distribution function approach*. Econometrica, 1996, **64**(3): 597–622.