Applied Mathematics, **2011, 2, 1309-1312**
doi:10.4236/am.2011.210182 Published Online October 2011 (http://www.SciRP.org/journal/am)

A Note on Change Point Detection Using Weighted Least Square

Reza Habibi
Department of Statistics, Central Bank of Iran, Tehran, Iran
E-mail: *rezahabibi*2681@*yahoo.com*
Received July 3, 2011; *revised July* 14, 2011; *accepted July* 22, 2011

Abstract

This paper is concerned with the application of weighted least square method in change point analysis. Testing shift in the mean normal observations with time varying variances as well as of a GARCH time series are considered. For both cases, the weighted estimators are given and their asymptotic behaviors are studied. It is also described that how the resampling methods like Monte Carlo and bootstrap may be applied to compute the finite sample behavior of estimators.

Keywords: Bootstrap, Brownian Bridge, Change Point, GARCH Series, Testing Shift, Monte Carlo, Weighted Least Square

1. Introduction

Change point analysis has been received considerable attentions in statistical literatures. This topic is studied from the frequentist and Bayesian point of view, by parametric and nonparametric approaches, with univariate and multivariate observations and in independent and correlated data. Three important references are Csorgo and Horvath [1], Chen and Gupta [2] and Khodadadi and Asgharian [3]. Bai [4] tested shift in mean of a linear process using the ordinary LS (OLS) approach. In many practical situations, however, it is advised to apply the weighted LS (WLS). In the current paper, we consider the WLS method for change point detection. The approach of derivation test statistics is similar to Bai [4], however, it is described briefly, as follows.

Suppose that $x_1,\cdots,x_n$ is a sequence of observations such that

$$x_i = E(x_i)+\varepsilon_i,\ i=1,2,\cdots,n,$$

where $E(x_i)=\mu_1$, for $i=1,\cdots,k_0$, and $=\mu_2$, for $i=k_0+1,\cdots,n$. Let w_i, $i=1,2,\cdots,n$ be non-negative numbers and called them weights. To make inference about μ_1, μ_2 and k_0, it is enough to minimize $\sum_{i=1}^{n} w_i\left(x_i-E(x_i)\right)^2$ with respect to μ_1, μ_2 and k_0. Note that

$$\sum_{i=1}^{n} w_i\left(x_i-E(x_i)\right)^2=\sum_{i=1}^{k_0} w_i\left(x_i-\mu_1\right)^2+\sum_{i=k_0+1}^{n} w_i\left(x_i-\mu_2\right)^2.$$

For a fixed $k=1,\cdots,n-1$, define

$$\hat{\mu}_{1k}=\left(\sum_{i=1}^{k} w_i x_i\right)\Big/\sum_{i=1}^{k} w_i,$$

and $\hat{\mu}_{2k}=\left(\sum_{i=k+1}^{n} w_i x_i\right)\Big/\sum_{i=1}^{k} w_i$. Following Bai [4] $\hat{k}$ (WLS estimation of k_0) is the minimizer of

$$\sum_{i=1}^{k} w_i\left(x_i-\hat{\mu}_{1\hat{k}}\right)^2+\sum_{i=k+1}^{n} w_i\left(x_i-\hat{\mu}_{2k}\right)^2.$$

One can see that $\hat{k}$ is the maximizer (argmax) of v_k given by

$$v_k=\sqrt{p_k\left(1-p_k\right)}\left|\hat{\mu}_{1\hat{k}}-\hat{\mu}_{2k}\right|,$$

at which $p_k=\sum_{i=1}^{k} w_i\Big/\sum_{i=1}^{n} w_i$. Therefore, the WLS estimate of μ_1 is $\hat{\mu}_1=\hat{\mu}_{1\hat{k}}$. To write $\hat{\mu}_2=\hat{\mu}_{2k}$, it is enough to replace $\sum_{i=1}^{\hat{k}}$ with $\sum_{i=\hat{k}+1}^{n}$ in $\hat{\mu}_1$. The large values of test statistic

$$T=\max_{1\le k\le n-1}\left|v_k\right|,$$

rejects the null hypothesis of no change point $H_0:\mu_1=\mu_2$ versus $H_1:\mu_1\ne\mu_2$. In practice, w_i's may be deterministic or random. They can be known or they may be function of unknown parameters (see the Example 2, as follows). In these cases, they are replaced by their estimations $\hat{w}_i$'s and consequently, p_k's are changed to

$\hat{p}_k$'s.

Under the null hypothesis, the plot of v_k against the number of observation k oscillates around zero. It remains between two specified boundaries (horizontal lines) with high probability. When there is a change in mean, the plot of $|v_k|$ creates a peak out of a boundary (see the following examples). Two horizontal lines (in examples) are obtained by the simulating null distribution using the Monte Carlo method. We can detect the change if T exceeds the boundary value at k_0. This suggests that the change point estimator is given by

$$\hat{k}_n = \arg\max_{1\le k\le n-1} |v_k|.$$

Remark 1. This problem also appears in continuous time processes cases. Suppose that S_t denote the time t (in continuous case) price of a specified stock. Let $U_t = \log(S_t)$. The Black Scholes formula implies that

$$dU_t = \gamma_t dt + \sigma_t dW_t,$$

where W_t is a Brownian motion over $t \in [0,T]$. Here, we assume that σ_t follows a GARCH(1,1) process and the mean process γ_t has a change point in t^*. Suppose that the process U_t is observed at $n+1$ equidistant discrete times $0 = t_0 < t_1 < \cdots < t_n = T$, with $t_i = i\Delta_n$ and $n\Delta_n = T$. For simplicity reasons, we write $U_i = U_{t_i}$, $W_i = W_{t_i}$, $\gamma_{t_i} = \gamma_i$ and $\sigma_{t_i} = \sigma_i$. Also, we assume that t^* is one of t_i's. The Euler approximation to the solution of above SDE is

$$U_i = U_{i-1} + \gamma_i \Delta_n + \sigma_i (W_i - W_{i-1}).$$

Let $V_i = \Delta_n^{-1}(U_i - U_{i-1})$ and $\eta_i = \Delta_n^{-1/2}(W_i - W_{i-1})$. Then $V_i = \gamma_i + \Delta_n^{-1/2}\sigma_i\eta_i$ with σ_i is a GARCH (1,1) defined in above.

Example 1. Shift in mean, time varying variances. Change point detection in the mean of normal observations is studied well, see Khodadadi and Asgharian [3]. An crucial assumption in this problem is fixing the variances after and before change point. Change point detection in variance of normal observations is another independent inferential problem. Change point detection in mean and variance at the same time is also studied. In this example, we consider the change point detection in mean when $\mathrm{var}(\varepsilon_i) = \sigma_i^2$, $i = 1, 2, \cdots, n$. Here, $w_i = 1/\sigma_i^2$ and $p_k = \sum_{i=1}^k \sigma_i^{-2} \Big/ \sum_{i=1}^n \sigma_i^{-2}$. The plots of $|v_k|$ under H_0 (**Figure 1**, page 6) and under H_1 (**Figure 2**, page 6) are given as follows. Here, n = 100, $k_0 = 30$, $\mu_1 = 0$, $\mu_2 = 2$ and $\sigma_i^2 = i/n$. The two horizontal lines are ± 3.1 obtained by a Monte Carlo with $R = 3000$ repetitions.

Remark 2. In above example, let $\sigma_i^2 = \sigma^2 = 1$. Then our procedure reduces to test statistic proposed by Bai [4]. By the way, we let $w_1, \cdots, w_n \overset{iid}{\sim} U(0,1)$. The two horizontal action lines are ± 2.575. The plot of v_k under H_1 (**Figure 3**, page 6) when $n = 100$, $\mu_1 = 0$ and $\mu_2 = 2$ shows that our method works well again.

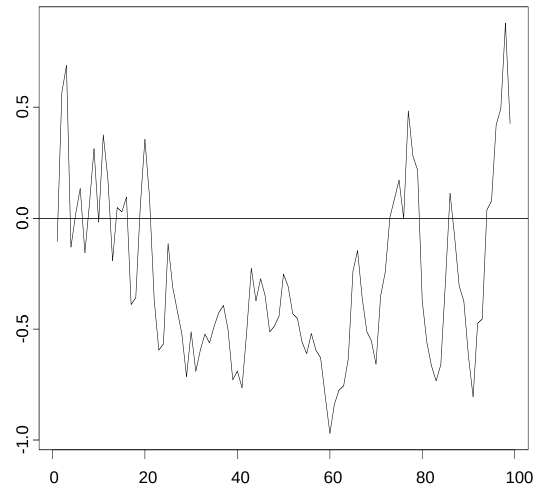

Figure 1. Time-varying-variance: H_0.

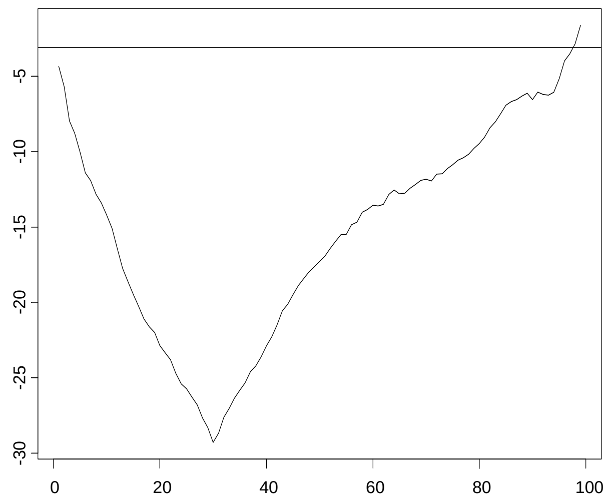

Figure 2. Time-varying-variance: H_1.

Example 2. Shift in mean, GARCH process. Lee *et al.* [5] studied change point analysis in regression models with ARCH errors. They used the maximum of CUSUM of square, based on estimated residuals, as test statistic. Here, we use the WLS method to change point detection in mean of GARCH process. Assume that ε_i's come from a GARCH(p,q) process, that is $\varepsilon_i = h_i\xi_i$, where ξ_i, $i = 1, 2, \cdots$ are *iid* random variables with zero mean and unit variance. The conditional variance h_i's are given by

$$h_i = \alpha_0 + \alpha(B)\varepsilon_i + \beta(B)h_i,$$

where $\alpha(B) = \alpha_1 B + \cdots + \alpha_q B^q$ and

$\beta(B) = \beta_1 B + \cdots + \beta_p B^p$ with $Bx_i = x_{i-1}$. If we want to apply the above mentioned method here, we should let $w_i = 1/h_i$. To see this, note that it is enough to minimize

$$\sum_{i=1}^{n} \xi_i^2 = \sum_{i=1}^{k_0} \left(1/h_i^2\right)\left(x_i - \mu_1\right)^2 + \sum_{i=k_0+1}^{n} \left(1/h_i^2\right)\left(x_i - \mu_2\right)^2,$$

with respect to unknown parameters. Therefore, $w_i = 1/h_i$. In practice, h_i^2's are unknown and they are replaced by their estimations $\hat{h}_i^2$. The two horizontal action lines are ± 2.876. The plot of v_k under H_1 (**Figure 4**, page 6) when $n = 1000$, $k_0 = 300$, $\mu_1 = 1$ and $\mu_2 = 3$ shows that our method works well again. The error process ε_i is GARCH(1,1) when $\alpha_0 = 0.01$, $\alpha_1 = 0.05$ and $\beta_1 = 0.9$. Since $\alpha_1 + \beta_1 < 1$ the GARCH series is stationary. In the next section, we study the application of bootstrap method in our problem.

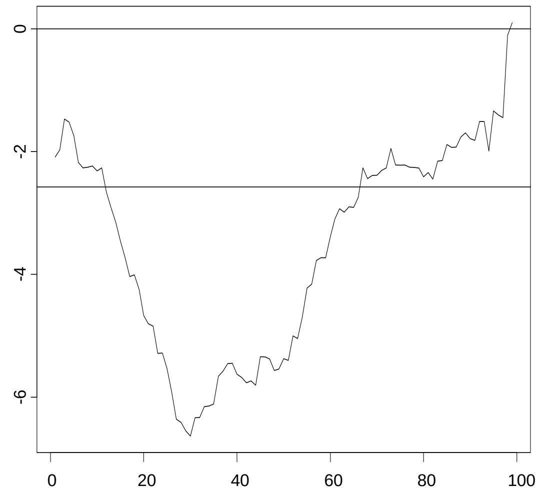

Figure 3. Constant variance: H_1.

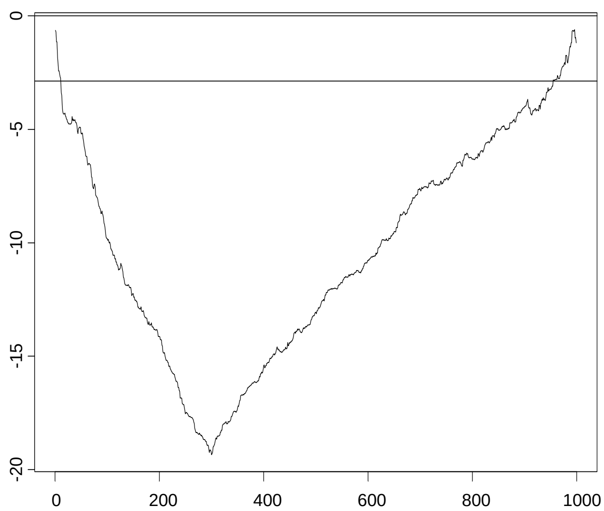

Figure 4. GARCH series: H_1.

2. Bootstrap Method

The WLS estimators appear, again, in bootstrap inference case. Bootstrap methods are strong practical solutions to the complicated problems. Chatterjee and Bose [6] proposed generalized bootstrap for estimating equations by imposing random weights (say multinomial weights for paired bootstrap) to the system of estimating equations. As stated by Chatterjee and Bose [6], this is equivalent to include the random weights to the original LS (or WLS) objective function. However, Chatterjee and Bose [6] didn't consider the change point version of their work. To extend work of Chatterjee and Bose [6] to the change point analysis, note that the bootstrapped WLS estimators of μ_1, μ_2 and k_0 are the minimizers of

$$\sum_{i=1}^{k_0} u_{ik_0} w_i \left(x_i - \mu_1\right)^2 + \sum_{i=k_0+1}^{n} u_{ik_0}^* w_i \left(x_i - \mu_2\right)^2.$$

The bootstrap estimator of change point ($\hat{k}_B$) is the minimizer of

$$\sum_{i=1}^{k_0} u_{ik_0} w_i \left(x_i - \hat{\mu}_{1k_0}^B\right)^2 + \sum_{i=k_0+1}^{n} u_{ik_0}^* w_i \left(x_i - \hat{\mu}_{2k_0}^B\right)^2,$$

over $k_0 = 1, \cdots, n-1$. In the current formula,

$$\hat{\mu}_{1k_0}^B = \left(\sum_{i=1}^{k_0} u_{ik_0} w_i x_i\right) \Big/ \sum_{i=1}^{k_0} u_{ik_0} w_i,$$

and $\hat{\mu}_{2k_0}^B = \left(\sum_{i=k_0+1}^{n} u_{ik_0}^* w_i x_i\right) \Big/ \sum_{i=k_0+1}^{n} u_{ik_0}^* w_i$. Finally,

$$\hat{\mu}_{iB} = \hat{\mu}_{i\hat{k}_B}^B, \; i = 1, 2.$$

In above formulas, random vectors $U = \left(u_{1k_0}, \cdots, u_{k_0 k_0}\right)$ and $U^* \left(u_{k_0+1,k_0}^*, \cdots, u_{nk_0}^*\right)$ specify the type of bootstrap method. For example for classical paired bootstrap has U multinomial distribution with parameters $\left(k_0, 1/k_0, \cdots, 1/k_0\right)$ and U^* is distributed as multinomial with parameters $\left(k_0^*, 1/k_0^*, \cdots, 1/k_0^*\right)$, at which $k_0^* = n - k_0$. Under the null hypothesis, since $k_0 = n$, it is enough to let $U = \left(U, U^*\right)$ has multinomial distribution with parameters $\left(n, 1/n, \cdots, 1/n\right)$. As it is stated by Chatterjee and Bose [6], the other bootstrap methods in the literatures like the Bayesian bootstrap, the deleted d-jackknives, and the bootstrap clone are also special cases of the above bootstrap formulation. By running the bootstrap method to data, and computing the above formula, one can derive the bootstrap quantile of weighted test statistic. Also, one can remove the bias of test statistic and construct confidence intervals based on bootstrap, we have done these calculations and it has been seen that the results are very good. Interested reader can refer to Habibi [7].

3. References

[1] M. Csorgo and L. Horvath, "Limit Theorems in Phange-Point Analysis," Wiley & Sons, New York, 1997.

[2] J. Chen and A. K. Gupta, "Parametric Statistical Change Point Analysis," Birkhäuser, Basel, 2000.

[3] A. Khodadadi and M. Asgharian, "Change Point Problem and Regression: An Annotated Bibliography," Technical Report, McGill University, Montreal, 2004.

[4] J. Bai, "Least Squares Estimation of a Shift in Linear Processes," *Journal of Time Series Analysis*, Vol. 15, No. 5, 1994, pp. 453-472. doi:10.1111/j.1467-9892.1994.tb00204.x

[5] S. Lee, Y. Tokutsu and K. Maekawa, "The CUSUM Test for Parameter Change in Regression Models with ARCH Errors," *Journal of Japan Statistical Society*, Vol. 3, 2004, pp. 173-186.

[6] S. Chatterjee and A. Bose, "Generalized Bootstrap for Estimating Equations," *Annals of Statistics*, Vol. 33, No. 1, 2005, pp. 414-436. doi:10.1214/009053604000000904

[7] R. Habibi, "Change Point Detection Using Weighted Least Square," Technical Report, Central Bank of Iran, 2010.

