Change Detection Experimentation for Time Series data by New Sequential Probability Ratio

Yoshihide Koyama *, **Tetsuo Hattori,**
Graduate School of Engineering, Kagawa University,
2217-20, Hayashi-cho Takamatsu, Kagawa 761-0396, Japan

Hiromichi Kawano
NTT Advanced Technology Corporation, 19-18, Nakamachi
Musashino, Tokyo 180-0006, Japan

Katsunori Takeda
Canon IT Solutions Inc., Tosabori Dai Bldg. 2-2-4, Tosabori,
Nishi-ku,Osaka-shi, Osaka 550-0001, Japan

Abstract

Previously, we have proposed a novel method using New Sequential Probability Ratio (NSPR) for the structural change detection problem of ongoing time series data instead of using SPRT (Sequential Probability Ratio Test). In this paper, for comparison, we present the experimental results by applying the both methods, i.e., NSPR and SPRT, to time series data that are generated by a multiple regression model in the case where one explanatory variation is a periodic function (sine function). And also we discuss the effectiveness of the both methods.

***Keywords*:** Time series data, structural change detection, SPRT, New Sequential Probability Ratio (NSPR)

1. Introduction

Change point detection (CPD) problem in time series (see Refs.1-7) is to identify whether the generation structure of monitoring data has changed at some time point by some reason, or not. We consider that the problem is very important and that it can be applied to a wide range of application fields.

For the structural change detection problem of ongoing time series data, we have already proposed some kinds of methods (see Refs. 6, 8-12). In this paper, we present the results of comparison experiment by two methods that are Sequential Probability Ratio Test (SPRT) and New Sequential Probability Ratio (NSPR) (Refs. 6, 11-12).

2. SPRT and Chow Test

2.1. *SPRT*

The Sequential Probability Ratio Test (SPRT) is used for testing a null hypothesis H_0 (e.g. the quality is under pre-specified limit 1%) against hypothesis H_1 (e.g. the quality is over pre-specified limit 1%). And it is defined as follows:

Let Z_1, Z_2, …Z_i be respectively observed time series data at each stage of successive events, the probability ratio λ_i is computed as follows.

*2217-20 Hayashi Cho Takamatsu City, Kagawa, 761-0396 Japan,
Kagawa University, Email: hattori@eng.kagawa-u.ac.jp

$$\lambda_i = \frac{P(Z_1 \mid H_1) \cdot P(Z_2 \mid H_1) \cdots P(Z_i \mid H_1)}{P(Z_1 \mid H_0) \cdot P(Z_2 \mid H_0) \cdots P(Z_i \mid H_0)} \quad (1)$$

where $P(Z \mid H_0)$ denotes the distribution of Z if H_0 is true, and similarly, $P(Z \mid H_1)$ denotes the distribution of Z if H_1 is true.

Two positive constants C_1 and C_2 ($C_1 < C_2$) are chosen. If $C_1 < \lambda_i < C_2$, the experiment is continued by taking an additional observation. If $C_2 < \lambda_i$, the process is terminated with the rejection of H_0 (acceptance of H_1). If λ_i < C1, then terminate this process with the acceptance of H_0.

$$C_1 = \frac{\beta}{1-\alpha}, \quad C_2 = \frac{1-\beta}{\alpha} \quad (2)$$

where α means type I error (reject a true null hypothesis), and β means type II error (accept a null hypothesis as true one when it is actually false).

2.2 *Procedure of SPRT*

The concrete procedure of applying the SPRT method to the structural change detection problem is described.

[Step 1] Make a prediction expression and set the tolerance band (***a***) (e.g. ***a***=2σ_s) that means permissible error margin between the predicted data and the observed one. (σ_s denotes a standard deviation in learning sample data at early stage.)

[Step 2] Set up the null hypothesis H_0 and alternative hypothesis H_1.

H_0 : Change has not occurred yet.

H_1 : Change has occurred.

Set the values α, β and compute C_1 and C_2, according to Eq. (2). Initialize i = 0, $\lambda_0 = 1$.

[Step 3] Incrementing i (i = i+1), observe the following data y_i. Evaluate the error | ε_i | between the data y_i and the predicted value from the aforementioned prediction expression.

[Step 4] Judge as to whether the data y_i goes in the tolerance band or not, i.e., the ε_i is less than (or equal to) the permissible error margin or not. If it is Yes, then set λ_i = 1 and return to Step3. Otherwise, advance to Step5.

[Step 5] Calculate the probability ratio λ_i, using the below Eq.(3) that is equivalent to Eq.(1).

$$\lambda_i = \lambda_{i-1} \frac{P(\varepsilon_i \mid H_1)}{P(\varepsilon_i \mid H_0)} \quad (3)$$

where, if the data y_i goes "OUT" from the tolerance band, $(P(\varepsilon_i \mid H_0), P(\varepsilon_i \mid H_1)) = (\theta_0, \theta_1)$, otherwise (i.e., the data y_i goes "IN"), $(P(\varepsilon_i \mid H_0), P(\varepsilon_i \mid H_1)) = ((1-\theta_0), (1-\theta_1))$.

[Step 6] Execution of testing.

(i) If the ratio λ_i is greater than C_2 (= $(1-\beta)/\alpha$), dismiss the null hypothesis H_0, and adopt the alternative hypothesis H_1, and then End.

(ii) Otherwise, if the ratio λ_i is less than C_1 (= $\beta/(1-\alpha)$), adopt the null hypothesis H_0, and dismiss the alternative hypothesis H_1, and then set λ_i = 1 and return to Step3.

(iii) Otherwise (in the case where $C_1 \le \lambda_i \le C_2$), advance to Step7.

[Step 7] Observe the following data y_i incrementing i. Evaluate the error | ε_i | and judge whether the data y_i goes in the tolerance band, or not. Then, return to Step5 (λ_i calculation).

2.3 *Chow Test*

The well-known Chow Test checks if there are significant differences or not, among residuals for three Regression Lines, where regression Line 1 obtained from the data before a change point tc, Line 2 from the data after tc, and Line 3 from the whole data so far, by setting up the change point hypothesis at each point in the whole data (Fig.1).

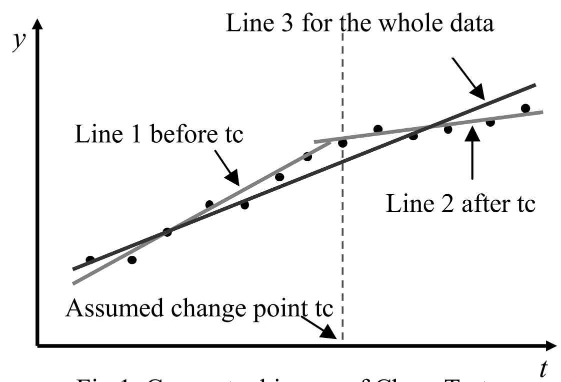


Fig.1. Conceptual image of Chow Test

3. New Sequential Probability Ratio (NSPR)

3.1. *Definition of NSPR*[11-12]

Let $C_n = a_1a_2,\ldots,a_i,\ldots a_n \quad a_i \in \{\text{IN,OUT}\}$ be a string (or symbol sequence) obtained from the observed data.
Let θ_i and $\tilde{\theta}_i$ be the conditional probability that receives the observed data a_i in the state $\boldsymbol{S_0}$ and $\boldsymbol{S_1}$, respectively, where $\boldsymbol{S_0}$ and $\boldsymbol{S_1}$ mean the state of unchanged state and changed state, respectively.
That is, it means that $\theta_i \in \{R, 1-R\}$ and $\tilde{\theta}_i \in \{R_c, 1-R_c\}$, respectively, where $R = P(\text{OUT} \mid \boldsymbol{S}_0)$, $R_c = P(\text{OUT} \mid \boldsymbol{S}_1)$.
And let $P(C_n, \text{H}_0)$ and $P(C_n, \text{H}_1)$ be the joint probability of the symbol sequence C_n that happen with the event H_0 (the structural change is not occurred) and H_1 (the change is occurred), respectively.
Then, the following equations hold.

$$P(a_1 \ldots a_n, \text{H}_0) = (1-\gamma)^n \theta_1 \ldots \theta_n = (1-\gamma)^n \prod_{i=1}^{n} \theta_i \tag{4}$$

$$\begin{aligned} P(a_1 \ldots a_n, \text{H}_1) &= \gamma \prod_{i=1}^{n} \tilde{\theta}_i + ((1-\gamma)\theta_1)(\gamma \prod_{i=2}^{n} \tilde{\theta}_i) \\ &\quad + ((1-\gamma)^2 \theta_1 \theta_2)(\gamma \prod_{i=3}^{n} \tilde{\theta}_i) + \ldots \\ &= \sum_{k=1}^{n} ((1-\gamma)^{k-1} \cdot \prod_{j=0}^{k-1} \theta_j)(\gamma \prod_{i=k}^{n} \tilde{\theta}_i) \end{aligned} \tag{5}$$

$$P(\text{H}_0 \mid a_1 \ldots a_n) \equiv P(\text{H}_0 \mid C_n) = \frac{P(C_n, \text{H}_0)}{P(C_n)} \tag{6}$$

$$P(\text{H}_1 \mid a_1 \ldots a_n) \equiv P(\text{H}_1 \mid C_n) = \frac{P(C_n, \text{H}_1)}{P(C_n)} \tag{7}$$

Using these above equations (Eq.(4)-(7)), the New Sequential Probability Ratio (NSPR) that we propose can be represented as follows.

$$\begin{aligned} \text{NSPR} &\equiv \frac{P(\text{H}_1 \mid a_1 \ldots a_n)}{P(\text{H}_0 \mid a_1 \ldots a_n)} = \frac{P(\text{H}_1 \mid C_n)}{P(\text{H}_0 \mid C_n)} = \frac{P(C_n, \text{H}_1)}{P(C_n, \text{H}_0)} \\ &= \frac{\sum_{k=1}^{n} ((1-\gamma)^{k-1} \cdot \prod_{j=0}^{k-1} \theta_j)(\gamma \prod_{i=k}^{n} \tilde{\theta}_i)}{(1-\gamma)^n \prod_{i=1}^{n} \theta_i} \end{aligned} \tag{8}$$

3.2. *Recursive Form*

When the NSPR is greater than 1.0, we can regard that the structural change has been occurred before the present time.
In the similar way as SPRT (see Eq.(3)), the definition of NSPR can also be described in a recursion formula. Let Λ_n be the value of NSPR defined by Eq.(8), we have the following recursive equation.

$$\Lambda_n = \frac{1}{1-\gamma}(\Lambda_{n-1} + \gamma)(\frac{\tilde{\theta}_n}{\theta_n}) \tag{9}$$

where $\Lambda_0 = 0,\ \theta_0 = 1,\ \tilde{\theta}_0 = 1$.

4. Experimentation for Comparison

We have done the comparison experimentation by SPRT and NSPR for a time series data that is generated based on the following equations.

$$\begin{cases} y = x_1 + 20x_2 + 5 + \varepsilon & (\text{t} \le \text{t}_\text{c}) \\ y = x_1 + 10x_2 + 5 + \varepsilon & (\text{t}_\text{c} \le \text{t}) \end{cases} \tag{10}$$

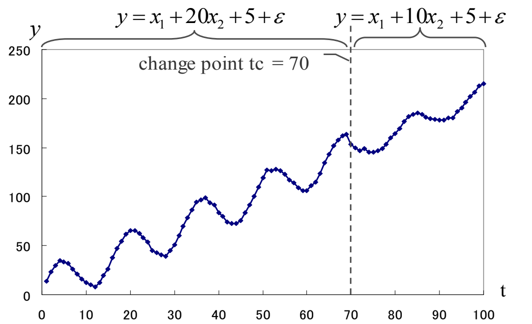
where $\varepsilon \sim N(0,\sigma^2)$, i.e., the error ε is subject to the Normal Distribution with the average 0 and the variation σ^2, and tc means the change point. In addition, we have set tc=70 and σ=5. The number of the time points is 100 (i.e. .
Setting the two kinds (or cases) of coefficients in Eq.(10) as shown in Eq.(11) for Case 1 and Eq.(12) for Case 2, respectively. And we have generated 200 sets of time series data for the two kinds (or cases) of time series, respectively.

$$x_1 = t,\ x_2 = \sin\left(\frac{1}{8}\pi t\right) \tag{11}$$

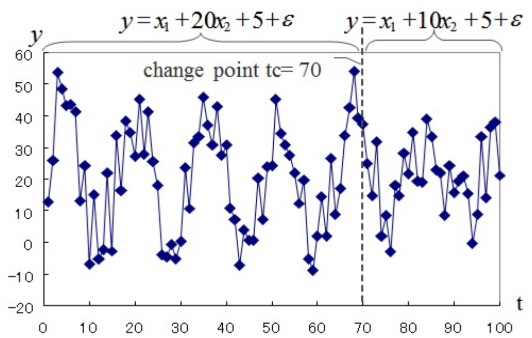
$$x_1 = random,\ x_2 = \sin\left(\frac{1}{8}\pi t\right) \tag{12}$$

For the two kinds of time series data (400 time series data in total), we have examined the change detection ability and the accuracy of each method (SPRT and NSPR) by setting some various evaluation criteria.
In the experimentation, the prediction model is constructed using the first 40 time points and the rest

of time points from t=41 to t=100 are used for change detection.



(a) (Case 1)



(b) (Case 2)

Fig.2. Example of two kinds of time series data generated based on Eq.(10).
(a)Case1 (x_1=*t*, and $x_2 = \sin(\pi t/8)$).
(b)Case2 (x_1=*random*, and $x_2 = \sin(\pi t/8)$).

We define a two dimensional vector ***I***=(I_1, I_2) as follows.

$$\boldsymbol{I} = (I_1, I_2) = \left(\sum_{t=41}^{t=70} fr(t) \times |t - tc|, \sum_{t=71}^{t=100} fr(t) \times |t - tc| \right) \qquad (13)$$

where tc=70, and $fr(t)$ means the detection frequency at time point t in 200 time series data.

We can consider that the detection ability and the accuracy is high as the absolute value of ***I*** and I_1 approach to 0.
As a result, in the sense of the abovementioned criteria, the NSPR method has shown that it is superior to SPRT for the both kinds of time series data used in the experimentation.

References

1. C. G. E. P. Box and G. M. Jenkins, *Time Series Analysis: Forecasting and Control* (Prentice Hall, 1976).
2. Peter J. Brockwell and Richard A. Davis, *Introduction to Time Series and Forecasting*, 2nd edn. (Springer; 2003)
3. C. Han, P. K. Willet and D. A. Abraham, Some methods to evaluate the performance of Page's test as used to detect transient signals, *IEEE Trans. Signal processing*. **47**(8) (1999) 2112-2127.
4. S. D. Blostein, Quickest detection of a time-varying change in distribution, *IEEE Trans. Information Theory*. **37**(4) (1991) 1116-1122.
5. A. Wald, *Sequential Analysis* (John Wiley & Sons, 1947).
6. H. Kawano, T. Hattori, and K. Nishimatsu, Structural Change Point Detection Method of Time Series Using Sequential Probability Ratio Test -Comparison with Chow Test in the ability of early detection- (in Japanese), *IEEJ Trans. EIS*. **128**(4) (2008) 583-592.
7. G. C. Chow, Tests of Equality Between Sets of Coefficients in Two Linear Regressions, *Econometrica*. **28**(3) (1960) 591-605.
8. K. Takeda, T. Hattori, T. Izumi and H. Kawano, Extended SPRT for Structural Change Detection of Time Series Based on Multiple Regression Mode, in *Proceedings of the 15th International Symposium on Artificial Life and Robotics*, (Feb. 2010), ISBN: 978-4-9902880-4-4, pp.755-758.
9. K. Takeda, T. Hattori, Tetsuya Izumi and H. Kawano, Extended SPRT for Structural Change Detection of Time Series Based on Multiple Regression Model, *International Journal of Artificial Life and Robotics* Springer, ISSN: 1433-5298, **15**(4), (2010) 417-420.
10. T. Hattori and H. Kawano, Change Detection Method of Time Series as an Optimal Stopping Problem -- Constructive Proof of Optimal Solution Theorem --, in *Proc. of 2011 International Conference on Biometrics and Kansei Engineering*, (2011), IEEE Computer Society, ISBN: 978-0-7695-4512-7, pp.100-105.
11. Yoshihide Koyama, Tetsuo Hattori, Hiromichi Kawano, Model Introduced SPRT for Structural Change Detection of Time Series (I), *in Proc. of The International Conference on Artificial Life and Robotics* (Jan. 2014), ISBN: 978-4-9902880-8-2, pp.252-255.
12. Yoshihide Koyama, Tetsuo Hattori, Hiromichi Kawano, Model Introduced SPRT for Structural Change Detection of Time Series (II) --- Kansei Channel and Bayes' Estimation ---, *in Proc. of The International Conference on Artificial Life and Robotics* (Jan. 2014), ISBN: 978-4-9902880-8-2, pp.256-259.