QUANTITATIVE METHODS IN ECONOMICS
Vol. XII, No. 1, 2011, pp. 87– 96

ON THE CHOICE OF PARAMETERS OF CHANGE-POINT DETECTION WITH APPLICATION TO STOCK EXCHANGE DATA

Stanisław Jaworski
Department of Econometrics and Statistics
Warsaw University of Life Sciences – SGGW
e-mail: stanislaw_jaworski@sggw.pl
Konrad Furmańczyk
Department of Applied Mathematics
Warsaw University of Life Sciences– SGGW
e-mail: konrad_furmanczyk@sggw.pl

Abstract: Our paper is devoted to the study of V-Box Chart method in a parametric model. This algorithm is proposed to be used in the change-point detection in a sequence of observations. The choice of parameters in such an algorithm is heuristic. In our paper we use the mini-max rule for this choice and we control the probability that no signal is given, when the process is out of control as well as the probability of false alarm. We apply this algorithm to the detection of a change in stock exchange data.

Key words: V-Box Chart, mini-max rule, normal distribution

INTRODUCTION

This paper examines a new method of detecting a change in a sequence of observations. The main issue is to design a monitoring procedure that detects the change as quickly as possible. This procedure can be applied in many practical problems relating to signal processing, quality control, finance, clinical medicine.

The literature of various aspects and procedures of the change-point detection problem can be found in [Poor and Hadjiliadis 2009], [Basseville and Nikiforov 1993], [Brodsky and Darkhovsky 2002], [Lai 1995].

In the theory of classical control chart it is assumed that we observe the realization of the model

$$Y_i = \theta 1(i \geq q) + \eta_i, \tag{1}$$

where (η_i) is a noise process, θ is a jump, $q > 0$ is an unknown change point and

$$1(p) = \begin{cases} 1, & \text{if } p \text{ is true} \\ 0, & \text{if } p \text{ is false} \end{cases}.$$

In the parametric approach it is commonly accepted that the noise process (η_i) is Gaussian or has any known distribution of continuous type. Popular detection schemes in this case are CUSUM algorithms and EWMA control charts (see [Lu and Reynolds 1999], [Ritov 1990]).

An interesting algorithm related to the nonparametric model of (1) has been presented in [Rafajłowicz et al 2010]. It was connected with the Vertical Box Control Chart (V-Box Chart). In our work we use the idea of V-Box Chart in the parametric case, together with the mini-max decision procedure. This approach is very similar to the problem of ε-comparison of means of two normal distributions (see [Jaworski and Zieliński 2004]).

The choice of the parameter of the algorithm in [Rafajłowicz et al 2010] is heuristic. We give a simple numerical procedure for choosing the parameters of V-Box Chart in parametric case. For examination of the proposed method we give some examples of detecting changes in stock exchange data.

THEORETICAL BACKGROUND

We consider the following statistical model

$$Y_i = \theta 1(i \geq N+1) + \eta_i, \text{ for } i = 1,\ldots,N+1,$$

where $\theta \in R$, N is the sample size of past observations, Y_{N+1} is a new observation and (η_i) is a noise process. We assume that (η_i) are i.i.d. with symmetric, continuous distribution. The problem is to take one of the following decisions:

$$d_1 : |\theta| \leq \varepsilon \text{ or } d_2 : |\theta| > \varepsilon,$$

for a given $\varepsilon > 0$. We take into account the following class of decision procedures:

$$d(\gamma, H) = \begin{cases} d_1, \text{ if } T(H) > \gamma N \\ d_2, \text{ if } T(H) \leq \gamma N \end{cases},$$

where $\gamma \in (0,1)$, $T(H) = \sum_{j=1}^{N} 1\left(Y_{N+1} - H \leq Y_j \leq Y_{N+1} + H\right)$ and $H > 0$. Notice that $2H$ is the vertical length of box chart and γ is the critical proportion of points in the box chart.

The idea of the optimal choice of parameters γ and H relates to the loss function L given in the Table 1 bellow.

Table 1. Loss function

$L(d_i,\theta)$	$\theta: \lvert\theta\rvert \le \varepsilon$	$\theta: \lvert\theta\rvert > \varepsilon$
d_1	0	b
d_2	a	0

where a, b are some positive values. For the fixed H we choose γ satisfying the mini-max condition:

$$\max_{\theta} E_{\theta} L(d(\gamma, H), \theta) = \min_{\gamma^*} \max_{\theta} E_{\theta} L\left(d\left(\gamma^*, H\right), \theta\right).$$

In the case when the marginal distributions of the noise (η_i) are symmetric the mini-max condition is equivalent to the following one

$$aP_{\varepsilon}\left(T(H) \le \gamma N\right) = bP_{\varepsilon}\left(T(H) > \gamma N\right). \tag{2}$$

To prove the above equality, it is enough to note that, for $\theta > 0$, $P_{\theta}\left(T(H) \le \gamma N\right) = P_{-\theta}\left(T(H) \le \gamma N\right)$ and the mapping $\theta \mapsto P_{\theta}(T(H) \le \gamma N)$ is increasing function of θ (for details see [Furmańczyk and Jaworski 2011]). Then (2) follows from the equation

$$\begin{aligned} E_{\theta} L\left(d\left(\gamma, H\right), \theta\right) = \ & aP_{\theta}\left(T(H) \le \gamma N\right) 1(-\varepsilon \le \theta \le \varepsilon) + \\ & + bP_{\theta}\left(T(H) > \gamma N\right)\left(1 - 1(-\varepsilon \le \theta \le \varepsilon)\right). \end{aligned}$$

The equation (2) is a base of numerical consideration of the problem of optimal choice of parameters. But in order to simplify it and to improve numerical stability this equation should be transformed. First we note, that

$$P_{\theta}\left(T(H) \le \gamma N \mid Y_{N+1} = y\right) = P_{\theta}\left(\sum_{j=1}^{N} 1\left(y - H \le Y_j \le y + H\right) \le \gamma N\right) = P_{\theta}\left(B(H, y) \le \gamma N\right),$$

where $B(H, y)$ is the binomial random variable with N trials and success probability $\pi(H, y) = F(y + H) - F(y - H)$ and F is the marginal distribution function of the noise process (η_i). Using well known property (see Zieliński [2009]), we obtain

$$P_{\theta}\left(T \le \gamma N \mid Y_{N+1} = y\right) = P_{\theta}\left(B(H, y) \le \gamma N\right) = B_{N(1-\gamma), \gamma N+1}\left(1 - \pi(H, y)\right),$$

where $B_{N(1-\gamma),\gamma N+1}$ denotes the cumulative distribution function of random variable with beta distribution with parameters $N(1-\gamma)$ and $\gamma N + 1$. Thus, the equality (2) can be written in the following form:

$$\int_R B_{N(1-\gamma),\gamma N+1}(1-\pi(H,y))f(y-\varepsilon)dy = \frac{b}{a+b}, \tag{3}$$

where f is the marginal density of the noise process (η_i).

Observe that the equation (3) might not have a solution for too small or too large value of H. The left hand side of the equation (3) is equal to

$E_\varepsilon P_0(T(H) \le \gamma N \mid Y_{N+1})$. It is clear that if $H \to \infty$, then $\pi(H,y) \to 1$, which implies that $E_\varepsilon P_0(T(H) \le \gamma N \mid Y_{N+1}) \to 0$. On the other side, if $H \to 0$, then $\pi(H,y) \to 0$ and $E_\varepsilon P_0(T(H) \le \gamma N \mid Y_{N+1}) \to 1$.

In the next part of our paper we present the numerical solution of the equation (3) in the case when the marginal distribution of the noise process is the standard normal and we discuss the optimal choice of parameters γ, H. The presented method can be transformed - under some technical assumptions - to any symmetric continuous distribution F of the noise process (see [Furmańczyk and Jaworski 2011]]).

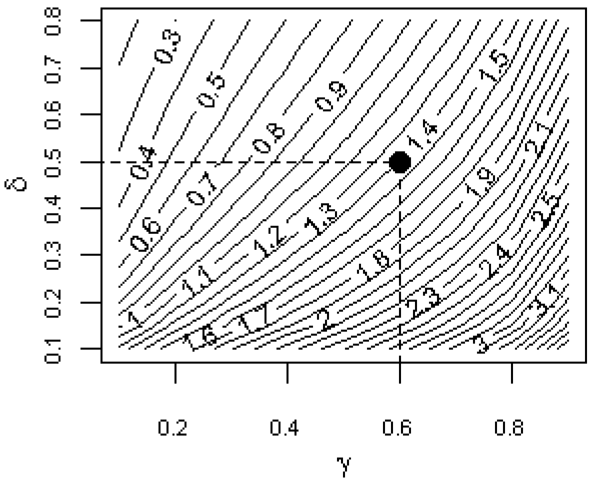
NUMERICAL CONSIDERATION

Let us denote $\delta = \frac{b}{a+b}$ and fix $\varepsilon = 1$. We want to solve the equation

$$\int_R B_{N(1-\gamma),\gamma N+1}(1-\pi(H,y))f(y-\varepsilon)dy = \delta,$$

with respect to γ and H. In order to avoid the problem of no solution (i.e., when H is too small or too large), we fix $\delta, \gamma \in (0,1)$ and solve the equation numerically with respect to H. We repeat this procedure for the following grid of parameters: $\delta, \gamma \in \{0.1 \cdot i : i = 1,2,\ldots,10\}$. Thus, we receive a set of solutions $H = H(\delta, \gamma)$, which is presented in Figure 1 as a contour plot.

Figure 1. Solutions of the mini-max equation



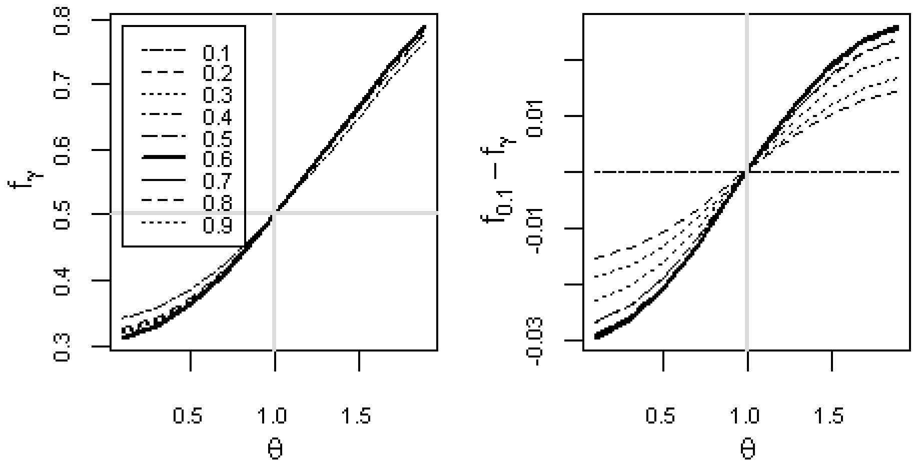
Source: own preparation

In Figure 1 the solution $H = H(0.5, 0.6) = 1.4$ is marked. The problem is the solution is not unique and we need the additional criterion to choose the best one. To solve it, we define the following function

$$f_\gamma : \theta \mapsto P_\theta\big(T(H(\gamma,\delta)) \le \gamma N\big) = \int_R B_{N(1-\gamma),\gamma N+1}\big(1-\pi\big(H(\gamma,\delta), y\big)\big)\phi(y-\theta)dy,$$

where ϕ is the density function of the standard normal distribution. Note that the function is symmetric. In Figure 2 the graph of f_γ for different values of γ and for the fixed $\delta = 0.5$ and $N = 20$ is presented (to notice the differences, we also put next the graph of $f_{0.1} - f_\gamma$).

Figure 2. Chart of functions: f_γ and $f_{0.1} - f_\gamma$.

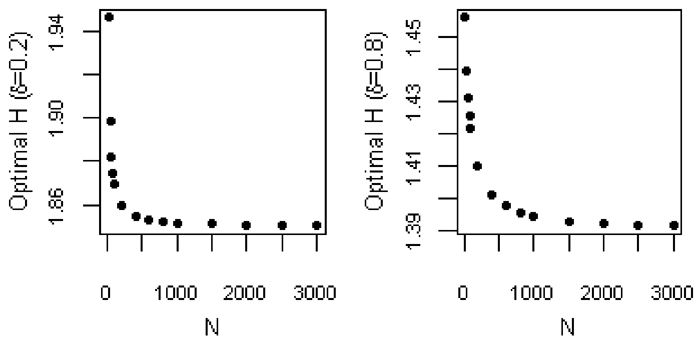


Source: own preparation

It is a natural requirement to have the probability $P_\theta(T(H(\gamma,\delta)) \leq \gamma N)$ (i.e., the probability of a decision d_2 : $|\theta| > \varepsilon$) as high as possible for $|\theta| > \varepsilon$, and as small as possible for $|\theta| \leq \varepsilon$. In this meaning, the optimal solution is obtained for $\gamma = 0.6$, which corresponds to $H = H(0.5, 0.6) = 1.4$.

Optimal solution for γ and H can be investigated asymptotically (i.e., when N tends to infinity). We consider $\delta = 0.2$ or $\delta = 0.8$. Figure 3 bellow presents the optimal H in our algorithm for a large sample N.

Figure 3. Asymptotic approach



Source: own preparation

From our simulation study we obtain that, when N increases, then the parameters H and γ are getting stabilized. For $\varepsilon = 1$ and $\delta = 0.8,$ we have $\gamma = 0.8$ and H is close to 1.3915. In the case when $\varepsilon = 1$ and $\delta = 0.2,$ we have $\gamma = 0.5$ and H is close to 1.8504.

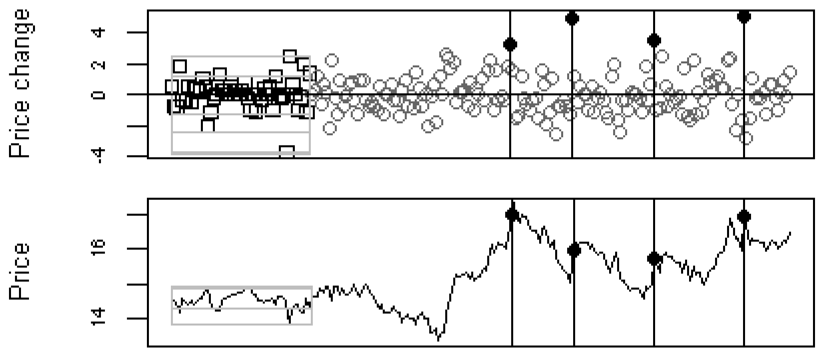
APLLICATION TO STOCK EXCHANGE DATA

In the case of the fixed N and ε the procedure can be represented by the value of $\delta = \frac{a}{a+b}$. In this chapter we apply our procedure for $N = 50$, $\varepsilon = 2$ and for three values of δ : 0.2, 0.5 and 0.8. We investigate the series of daily closing prices of Polish DTH platform "Cyfrowy Polsat" from 2010.06.25 until 2011.05.10. The value $\delta = 0.2$ relates to the high loss of making the wrong decision that the expected price change is not greater than the chosen level ε. It means we do not care for relatively small and medium price variability. Small and medium changes are more frequent than the large ones. Thus, the procedure is conservative and it does not reflect frequent price changes except for the largest ones. The case of $\delta = 0.8$ corresponds to the procedure sensitive for large changes, as we have high loss of making the wrong decision that the expected price change does not exceed or is less than the predefined level ε.

The value $\delta = 0.5$ expresses the strategy of choosing the more probable case: the new expected price level changed more than ε or not. The loss is equally balanced between the two possible decisions.

The three values of δ relate to different aspects of the time series prices. In the following figures we put two graphs. The first one shows the price changes in time and the second one presents the price levels. The shaded areas in the graphs represent the teaching part of the sample. The teaching observations are fixed. These observations establish a base to investigate each and next observation by the procedure. It means, we apply this procedure sequentially. Thus, the extreme changes are discovered and marked as black bullets in the charts. There are in Figure 4 a few black points. It is worthy to note that positive large changes are followed by the decreasing prices. If it were a rule we would have an excellent investment strategy for the considered time series.

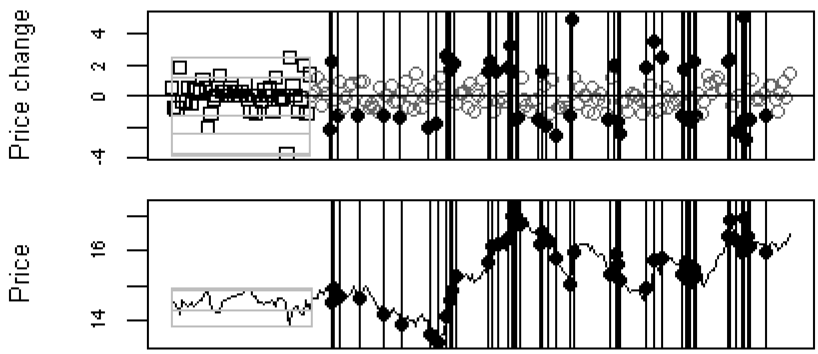
Figure 4. An application of the procedure ($\delta = 0.2,\ N = 50,\ H = 2.88$ and $\gamma = 0.5$)



Source: own preparation

Figure 5 represents the case when $\delta = 0.8$. Extreme points are much more exposed. After the teaching part of the sample we have the series of negative changes. Vertical lines cross the extreme points. Note that their compact allocation announce the price level change.

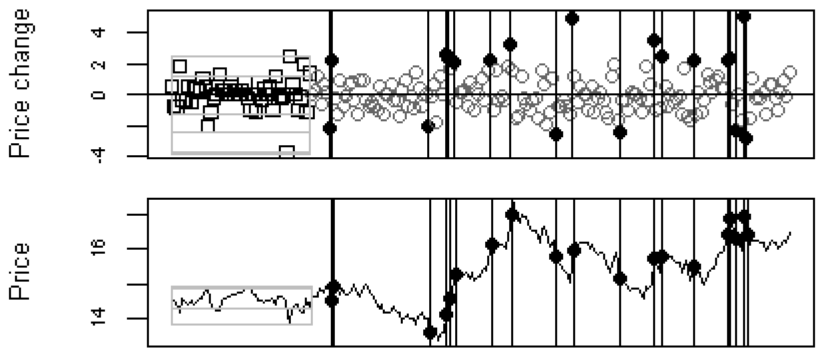
Figure 5. An application of the procedure ($\delta = 0.8,\ N = 50,\ H = 1.44$ and $\gamma = 0.6$)



Source: own preparation

Figure 6 represents the case when $\delta = 0.5$. The black points on the first chart of Figure 6 seem more likely to belong to the distribution with the expected value more different from zero than ε .

Figure 6. An application of the procedure ($\delta = 0.5,\ N = 50,\ H = 2.03$ and $\gamma = 0.5$)



Source: own preparation

In order to apply this procedure we transformed the time series of price changes. The values of the series were divided by the standard deviation of the teaching part of the sample. Simulation study shows that the distribution of the statistic T does not change significantly.

Table 2 presents the results of Monte–Carlo estimation of the probability $P_{\varepsilon}\left(T(H) \le \gamma N\right)$, where $N = 50$ and $\varepsilon = 2$:

Table 2. Monte-Carlo estimation of $P_{\varepsilon}\left(T(H) \le \gamma N\right)$; $N = 50$, $\varepsilon = 2$; for 100000 of replications

(δ, γ, H)	Not transformed series	Transformed series
(0.8, 0.6, 1.4404)	0.80009	0.79907
(0.2, 0.5, 2.8706)	0.20033	0.21194
(0.5, 0.5, 2.0002)	0.49832	0.50203

Source: own calculations

SUMMARY

Our paper is devoted to the construction a new algorithm assigned to the change-point detection in a sequence of observations. The choice of parameters in such an algorithm is based on the mini-max rule. Thus, we control both the probability that no signal is given when the process is out of control and the probability of false alarm. We apply this algorithm to detect the change in the stock exchange data. Due to the theoretical assumptions of the procedure, the empirical data must be transformed. Monte Carlo simulation shows that the transformation we used in the application part of the paper does not change the

theoretical results. The utility of the proposed algorithm depends on the specific features of the data. We gave an example of the series producing sudden changes in single observations with reference to its mean level change. This type of variability may be a rule in some time series, but it needs careful consideration. The proposed algorithm gives no indication on the causes from those such behavior occurs, but it responds quickly to large isolated jumps.

REFERENCES

Basseville M. and Nikiforov I. V. (1993) Detection of Abrupt Changes: Theory and Application. Upper Saddle River, NJ: Prentice-Hall.

Brodsky B. E. and Darkhovsky B. S. (2002) Nonparametric Methods in Change-Point Problems. Dordrecht, The Netherlands: Kluwer.

Furmańczyk K. and Jaworski S.. (2011) on internet http://p117650.users.sggw.pl/VBox/chart_box.pdf

Jaworski S., Zieliński W. (2004) A procedure for epsilon-comparison of means of two normal distribution, Applicationes Mathematicae, 31(2), pp. 155-160

Lai T. L. (1995) Sequential change-point detection in quality control and dynamical systems, J. Roy. Statist. Soc. Ser. B, vol. 57, pp. 613–658.

Lu C. W. and Reynolds M. R (1999) EMWA control charts for monitoring the mean of autocorrelated process, J. Quality Technol., vol. 3, pp. 166–188.

Poor H. V. and Hadjiliadis O. (2009) Quickest Detection, Cambridge, U.K. Cambridge Univ. Press.

Rafajłowicz E., Pawlak M., Steland A. (2010) Nonparametric sequential change-point detection by a vertically trimmed box metod, IEEE Transactions on Information Theory, 56(7), pp. 3621-3634.

Ritov Y. (1990) Decision theoretic optimality of the CUSUM procedure, Ann. Statist., vol. 18, pp. 1464–1469.

Zieliński W. (2009) A nonparametric confidence interval for At-Risk-of-Poverty-Rate, Statistics in Transition new series, 10(3), pp. 437-444.