Far East Journal of Theoretical Statistics
© 2016 Pushpa Publishing House, Allahabad, India
Published Online: April 2016
http://dx.doi.org/10.17654/TS052020075
Volume 52, Number 2, 2016, Pages 75-87 **ISSN: 0972-0863**

JACKKNIFE EMPIRICAL LIKELIHOOD TEST FOR CHANGES IN MEAN AND VARIANCE

Ying-Ju Chen[a], Wei Ning[b,*] and Arjun K. Gupta[b]

[a]Department of Information Systems and Analytics
Miami University
Oxford, OH 45056
U. S. A.

[b]Department of Mathematics and Statistics
Bowling Green State University
Bowling Green, OH 43403
U. S. A.
e-mail: wning@bgsu.edu

Abstract

In this paper, we propose a detection procedure based on the jackknife empirical likelihood for the changes in mean and variance. Simulations have been conducted to show the performance of the detecting procedure. Two real examples are also given: the stock market data and the air traffic data in Hsu [6] to illustrate the detecting procedure.

1. Introduction

Change-point problems have intrigued statisticians since Page [10, 11].

Received: October 24, 2015; Accepted: December 5, 2015
2010 Mathematics Subject Classification: 62G05, 62G10, 62G35.
Keywords and phrases: change-point, jackknife empirical likelihood, *U*-statistic.
*Corresponding author
Communicated by K. K. Azad

It is mainly because change-point problem can be applied in many areas, such as biology, economics, finance, geology, medicine, and so on. The commonly used approaches for change-point inference in the literature are the maximum likelihood ratio test, Bayesian information criterion, nonparametric test, sequential tests, etc. In general, the change-point problem includes hypothesis testing for the existence of change-point and the estimations of the number and location(s) of change-points.

Many scholars have studied on change-point problems in the past decades. For example, Hawkins [4] derived the likelihood ratio test and the asymptotic distribution of the test statistic under null hypothesis for detecting the mean change. Hsu [5] adopted the maximum likelihood to study the time of changes and variances at the different periods of the time series. Chen and Gupta [1] used the information approach to study the mean and variance change problem for univariate Gaussian model. Recently, Zou et al. [15] proposed a procedure based on the empirical likelihood by Owen [8] to detect changes in distributions, and showed the asymptotic distribution of the test statistic is the extreme value distribution. An extensive literature review of parametric and nonparametric change-point problems is provided by Csörgő and Horváth [3] and Chen and Gupta [2]. Readers are also referred to Owen [9] for more details of the empirical likelihood method.

In our work, we adopt the idea of one sample *U*-statistic from Jing et al. [7] and propose a nonparametric test for testing change-points in mean and variance. This article is structured as follows. In Section 2, we develop the JEL-type statistic based on *U*-statistic of mean and variance. The performance of the proposed JEL ratio test is studied through simulations in Section 3. We apply the proposed method to two real data sets: the stock market data and the air traffic data in Hsu [6] to show the testing procedure in Section 4. Discussion is provided in Section 5.

2. Methodology

Let $X_1, X_2, \ldots, X_n$ be independent random variables with means $\mathbb{E}X_i = \mu_i$ and variances $var(X_i) = \sigma_i^2$ for $i = 1, 2, \ldots, n.$ We are interested in testing

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_n = \mu$$

and

$$\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_n^2 = \sigma^2 \quad (\mu, \sigma^2 \text{ are unknown}) \tag{2.1}$$

versus the alternative

$$H_A : \mu_1 = \cdots = \mu_{k_1} \neq \mu_{k_1+1} = \cdots = \mu_{k_2} \neq \cdots \neq \mu_{k_q+1} = \cdots = \mu_n$$

and

$$\sigma_1^2 = \cdots = \sigma_{k_1}^2 \neq \sigma_{k_1+1}^2 = \cdots = \sigma_{k_2}^2 \neq \cdots \neq \sigma_{k_q+1}^2 = \cdots = \sigma_n^2. \tag{2.2}$$

Since the binary segmentation procedure (Vostrikova [14]) can be applied to detect multiple change-points, it is sufficient to study the null hypothesis (2.1) versus the alternative:

$$H_A : \mu_1 = \cdots = \mu_k \neq \mu_{k+1} = \cdots = \mu_n$$

and

$$\sigma_1^2 = \cdots = \sigma_k^2 \neq \sigma_{k+1}^2 = \cdots = \sigma_n^2. \tag{2.3}$$

Consider the kernel function $h : \mathbb{R}^2 \to \mathbb{R}^2$, where

$$h(X_i, X_j) = \begin{pmatrix} \frac{1}{2}(X_i + X_j) \\ \frac{1}{2}(X_i - X_j)^2 \end{pmatrix}. \tag{2.4}$$

Then we have

$$\mathbb{E}h(X_i, X_j) = \begin{pmatrix} \frac{1}{2}(\mu_i + \mu_j) \\ \frac{1}{2}(\sigma_i^2 + \sigma_j^2 + (\mu_i - \mu_j)^2) \end{pmatrix} = \begin{pmatrix} \mu \\ \sigma^2 \end{pmatrix} \text{ under } H_0.$$

A one sample U-statistic of degree 2 with the kernel function (2.4) is given by

$$U_n(X_1, X_2, \ldots, X_n) = \sum_{1 \le i_1 < i_2 \le n} \binom{n}{2}^{-1} h(X_{i1}, X_{i2}), \tag{2.5}$$

and

$$\mathbb{E}U_n = \begin{pmatrix} \frac{1}{n}\sum_{i=1}^{n} \mu_i \\ \frac{1}{n}\sum_{i=1}^{n} \sigma_i^2 + \frac{1}{n(n-1)} \sum_{1 \le i < j \le n} (\mu_i - \mu_j)^2 \end{pmatrix} = \begin{pmatrix} \mu \\ \sigma^2 \end{pmatrix} \text{ under } H_0.$$

Similar to Jing et al. [7], we apply the JEL method to the U_n. The jackknife pseudo-values are

$$\hat{V}_i = nU_n - (n-1)U_n^{(-i)}, \tag{2.6}$$

where $U_n^{(-i)} := U_n(X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n)$. One can check that

$$\mathbb{E}V_i = \begin{pmatrix} \mu_i \\ \sigma_i^2 + \mu_i^2 - \frac{2}{n-1} \sum_{1 \le k < l \le n} \mu_k \mu_l + \frac{2}{n-2} \sum_{\substack{1 \le k < l \le n \\ k, l \ne i}} \mu_k \mu_l \end{pmatrix}$$

for $i = 1, 2, \ldots, n$. Thus, we can apply the empirical likelihood approach to the $\hat{V}_i$'s based on the fact that $\hat{V}_i$'s are asymptotically independent under mild conditions (Shi [13]).

For the simplicity, we denote $\theta_{nk} = k/n$ and $\theta_{n0} = k^*/n$, where k^* is the true change location if it occurs. Throughout the whole article, we assume that $\theta_{n0} \to \theta_0 \in (0, 1)$ as $n \to \infty$. Let $p = (p_1, p_2, \ldots, p_{n\theta_{nk}})$ and $q = (q_{n\theta_{nk}+1}, q_{n\theta_{nk}+2}, \ldots, n)$ be probability vectors. If there exists a change at $k = k^*$, then we reject H_0 for small

$$L(\theta_{nk}) = \frac{\begin{array}{c}\sup\left\{\prod_{i=1}^{n\theta_{nk}} p_i \prod_{j=n\theta_{nk}+1}^{n} q_j \,\middle|\, \sum_{i=1}^{\theta_{nk}} p_i V_i = \begin{pmatrix}\mu_1\\ \sigma_1^2\end{pmatrix},\right.\\ \left.\sum_{j=n\theta_{nk}+1}^{n} q_j V_j = \begin{pmatrix}\mu_n\\ \sigma_n^2\end{pmatrix},\ \mu_1 = \mu_n = \mu,\ \sigma_1^2 = \sigma_2^2\right\}\end{array}}{\begin{array}{c}\sup\left\{\prod_{i=1}^{n\theta_{nk}} p_i \prod_{j=n\theta_{nk}+1}^{n} q_j \,\middle|\, \sum_{i=1}^{\theta_{nk}} p_i V_i = \begin{pmatrix}\mu_1\\ \sigma_1^2\end{pmatrix},\right.\\ \left.\sum_{j=n\theta_{nk}+1}^{n} q_j V_j = \begin{pmatrix}\mu_n\\ \sigma_n^2\end{pmatrix}\right\}\end{array}}.$$

Since $\prod_{i=1}^{n\theta_{nk}} p_i$ subject to $\sum_{i=1}^{n\theta_{nk}} p_i = 1$ attains its maximum $(n\theta_{nk})^{-n\theta_{nk}}$ at $p_i = (n\theta_{nk})^{-1}$ and $\prod_{j=n\theta_{nk}+1}^{n} q_j$ subject to $\sum_{j=n\theta_{nk}+1}^{n} q_j = 1$ attains its maximum $(n(1-\theta_{nk}))^{-n(1-\theta_{nk})}$ at $q_j = (n(1-\theta_{nk}))^{-1}$, the corresponding jackknife empirical log-likelihood ratio is obtained by

$$\begin{aligned} l(\theta_{nk}) &= -2\ln L(\theta_{nk}) \\ &= -2\sup\left\{\sum_{i=1}^{n\theta_{nk}} \ln(n\theta_{nk} p_i) + \sum_{j=n\theta_{nk}+1}^{n} \ln[n(1-\theta_{nk}) q_j] \mid \mu_1 = \right. \\ &\qquad \left. = \mu_n = \mu,\ \sigma_1^2 = \sigma_n^2 = \sigma^2 \right\}. \end{aligned} \tag{2.7}$$

We apply the Lagrange multiplier method in the following. Let

$$\begin{aligned} & f(p, q, \lambda_1, \lambda_2, \gamma_1, \gamma_2, \mu, \sigma^2) \\ &= \sum_{i=1}^{n\theta_{nk}} \ln p_i + \sum_{j=n\theta nk+1}^{n} \ln q_j - \gamma_1\left(\sum_{i=1}^{n\theta_{nk}} p_i - 1\right) - \gamma_2\left(\sum_{j=n\theta_{nk}+1}^{n} q_j - 1\right) \end{aligned}$$

$$- n\lambda_1^T\left(\sum_{i=1}^{\theta_{nk}} p_i V_i - \begin{pmatrix}\mu_1\\ \sigma_1^2\end{pmatrix}\right) - n\lambda_2^T\left(\sum_{j=n\theta_{nk}+1}^{n} q_j V_j - \begin{pmatrix}\mu_n\\ \sigma_n^2\end{pmatrix}\right), \tag{2.8}$$

where $\lambda_1^T = (\lambda_{11}, \lambda_{12})$ and $\lambda_2^T = (\lambda_{21}, \lambda_{22})$.

$$\frac{\partial f}{\partial p_i} = \frac{1}{p_i} - \gamma_1 - n\lambda_1^T V_i = 0 \Rightarrow \gamma_1 = n\theta_{nk} - n\lambda_1^T\begin{pmatrix}\mu\\ \sigma^2\end{pmatrix},$$

$$\frac{\partial f}{\partial q_j} = \frac{1}{q_j} - \gamma_2 - n\lambda_2^T V_j = 0 \Rightarrow \gamma_2 = n(1-\theta_{nk}) - n\lambda_2^T\begin{pmatrix}\mu\\ \sigma^2\end{pmatrix}.$$

The relationship between λ_1 and λ_2 can be obtained from the following derivatives:

$$\frac{\partial f}{\partial \mu} = -n\lambda_{11} - n\lambda_{21} = 0,$$

$$\frac{\partial f}{\partial \sigma^2} = -n\lambda_{12} - n\lambda_{22} = 0.$$

Thus, $\lambda_2 = -\lambda_1$. Consequently,

$$p_i = \frac{1}{\gamma_1 + n\lambda_1^T V_i} = \frac{1}{n\theta_{nk}} \frac{1}{1 + \frac{\lambda^T}{\theta_{nk}}(V_i - (\mu, \sigma^2)^T)},$$

$$q_j = \frac{1}{\gamma_2 + n\lambda_2^T V_j} = \frac{1}{n(1-\theta_{nk})} \frac{1}{1 - \frac{\lambda^T}{(1-\theta_{nk})}(V_j - (\mu, \sigma^2)^T)}.$$

Let $\gamma = -\lambda^T(\mu, \sigma^2)^T$. Then (2.7) can be simplified as

$l(\theta_{nk}, \lambda, \gamma)$

$$= 2\left\{\sum_{i=1}^{n\theta_{nk}} \ln(1 + \theta_{nk}^{-1}(\lambda^T V_i + \gamma)) + \sum_{j=n\theta_{nk}+1}^{n} \ln(1 - (1-\theta_{nk})^{-1}(\lambda^T V_j + \gamma))\right\}. \tag{2.9}$$

Define the score functions

$$J_1(\theta_{nk}, \lambda, \gamma) = \frac{\partial l(\theta_{nk}, \lambda, \gamma)}{2\partial\lambda} = \sum_{i=1}^{n\theta_{nk}} \frac{V_i}{\theta_{nk} + (\lambda^T V_i + \gamma)}$$

$$- \sum_{j=n\theta_{nk}+1}^{n} \frac{V_j}{1 - \theta_{nk} - (\lambda^T V_j + \gamma)},$$

$$J_2(\theta_{nk}, \lambda, \gamma) = \frac{\partial l(\theta_{nk}, \lambda, \gamma)}{2\partial\gamma} = \sum_{i=1}^{n\theta_{nk}} \frac{1}{\theta_{nk} + (\lambda^T V_i + \gamma)}$$

$$- \sum_{j=n\theta_{nk}+1}^{n} \frac{1}{1 - \theta_{nk} - (\lambda^T V_j + \gamma)}. \tag{2.10}$$

Then the maximal likelihood estimators $(\hat{\lambda}(\theta_{nk}), \hat{\gamma}(\theta_{nk}))$ of (λ, γ) can be obtained by solving

$$J_1(\theta_{nk}, \lambda, \gamma) = 0 \quad \text{and} \quad J_2(\theta_{nk}, \lambda, \gamma) = 0.$$

Therefore, we have $l(\theta_{nk}) = l(\theta_{nk}, \hat{\sigma}^2(\theta_{nk}), \hat{\lambda}(\theta_{nk}))$. The jackknife empirical log-likelihood ratio statistic is defined by

$$Z_n = \max_{\theta_{nk} \in \Theta_n} \{l(\theta_{nk})\},$$

where $\Theta_n = \{k/n, k = 1, 2, ..., n\}$. We reject H_0 for sufficient large value of Z_n. Zou et al. [15] suggested the use of the so-called "trimmed" likelihood ratio

$$Z'_n = \max_{\theta_{nk} \in \Theta'_n} \{l(\theta_{nk})\}$$

instead of Z_n, where $\Theta'_n = \{k/n, k = n_{T_1}, n_{T_1} + 1, ..., n - n_{T_2}\}$ since the empirical likelihood estimators of (λ, γ) may not exist when k or $n - k$ is too small, such as $n\theta_{nk} \sim O(1)$. Moreover, Perron and Vogelsane [12] mentioned that the choices of n_{T_1} and n_{T_2} might be rather arbitrary. We choose $n_{T_1} = n_{T_2} = 2\lfloor\sqrt{n}\rfloor$ for convenience, where $\lfloor x \rfloor$ means the largest integer not greater than *x*.

If we assume that there is exactly one change-point, then the maximum likelihood estimator for the change location k^* is given by

$$\hat{k}^* = n \min\{\theta_{nk} : Z'_n = l(\theta_{nk}),\ \theta \in \Theta'_{nk}\}. \tag{2.11}$$

3. Simulation Study

This section presents a simulation study to illustrate the performance of the proposed JEL method of detecting changes in mean and variance. Monte Carlo experiments with 1000 repetitions are implemented on normal, exponential, and Poisson distributions under various values of mean and variance with sample sizes 70, 100 and 150. The critical values of the test statistic for different distributions are obtained through simulations. The results corresponding to sample sizes 70, 100 and 150 are presented in Tables 1, 2 and 3. Corresponding standard errors are reported in parentheses.

Table 1. Power of JEL approach of change-point models in mean and variance with sample size 70 and $\alpha = 0.05$

	$N(0, 1),\ N(0,\ 2)$	$Exp(1),\ Exp(2)$	$Poisson(1),\ Poisson(2)$
k^*			
0	0.05	0.05	0.05
25	0.604 (0.0155)	0.238 (0.0135)	0.426 (0.0156)
35	0.697 (0.0145)	0.314 (0.0147)	0.489 (0.0158)
50	0.530 (0.0158)	0.243 (0.0136)	0.365 (0.0152)

* *Exp*(λ) where λ is the rate parameter

Table 2. Power of JEL approach of change-point models in mean and variance with sample size 100 and $\alpha = 0.05$

	$N(0, 1)$, $N(0, 2)$	$Exp(1)$, $Exp(2)$	$Poisson(1)$, $Poisson(2)$
k^*			
0	0.05	0.05	0.05
25	0.703 (0.0144)	0.301 (0.0145)	0.477 (0.0158)
50	0.890 (0.0099)	0.425 (0.0156)	0.620 (0.0153)
70	0.797 (0.0127)	0.308 (0.0146)	0.511 (0.0158)

*$Exp(\lambda)$ where λ is the rate parameter

Table 3. Power of JEL approach of change-point models in mean and variance with sample size 150 and $\alpha = 0.05$

	$N(0, 1)$, $N(0, 2)$	$Exp(1)$, $Exp(2)$	$Poisson(1)$, $Poisson(2)$
k^*			
0	0.05	0.05	0.05
40	0.920 (0.0086)	0.401 (0.0155)	0.618 (0.0154)
75	0.979 (0.0045)	0.454 (0.0157)	0.753 (0.0136)
120	0.882 (0.0102)	0.268 (0.0140)	0.494 (0.0158)

*$Exp(\lambda)$ where λ is the rate parameter

According to our simulation results, there is no surprise that the power increases as the sample size increases. In addition, when the normality assumption holds, the power performance of the test is much better than the power under other distribution settings. When the change location occurs in the middle of the sample, the power is higher than the power when the change occurs in other locations.

4. Application

In this section, we apply the testing procedure to two real data sets: the stock market data: the weekly closing values of Dow-Jones Industrial

Average from July 1, 1971 to August 2, 1974 and the air traffic data: the aircraft arrival times from a low altitude transitional control section near Newark Airport for the period from noon to 8:00 pm on April 30, 1969. Both data sets are from Hsu [6]. In order to find the potential change location in a wide range, we compute $Z'_n = \max_{\theta_{nk} \in \Theta'_n} \{l(\theta_{nk})\}$, where

$$\Theta'_n = \{k/n,\ k = n_T,\ n_T + 1,\ \ldots,\ n - n_T\} \text{ and } n_T = 2\lfloor \log(n) \rfloor.$$

4.1. Application to the stock market data

In this section, we study an application of JEL based detecting procedure to search potential changes in mean and variance for stock market price data in Hsu [6]. Let P_t be the stock price where $t = 1, 2, \ldots, 161.$ We transform the data into $R_t = (P_{t+1} - P_t)/P_t,\ t = 1, 2, \ldots, 161.$ The scatter plot of the return series $\{R_t\}$ is given in Figure 1. From Hsu [5], $\{R_t\}$ is a sequence of independent normal random variables with mean 0. Using the binary segmentation procedure along with the JEL based procedure, we are able to detect the potential changes in the $\{R_t\}.$ On the basis of our computations, at the first stage $Z'_n = \max_{\theta_{nk} \in \Theta'_n} \{l(\theta_{nk})\} = l(89) = 31.88398$ while the empirical critical value is only 20.34578 with the significance level $\alpha = 0.10.$ Transferring back to the price P_t, $t = 90$ is the location of change-point which is referred to the week of March 19 to March 23, 1973. Our conclusion matches Hsu's [5, 6] and Chen and Gupta [2] conclusions at this point.

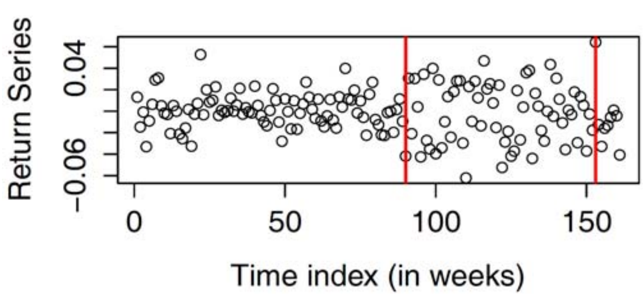

Figure 1. Return series $\{R_t\}$ of the weekly stock prices from 1971 to 1974.

Additionally, we continue to test the two sequences: R_t, $t = 1, \ldots, 89$ and R_t, $t = 90, \ldots, 161$. Our computational results show that there are no further changes in the sequence R_t, $t = 1, \ldots, 89$, but there is one more change at $t = 153$ in the sequence R_t, $t = 90, \ldots, 161$ which corresponds to the stock price P_{154} in the week of June 3 to June 7, 1974. The test statistic $Z'_n = l(153) = 64.18358 > 49.71162$, where 49.71162 is the empirical critical value of 90% confidence level. This finding is different from Chen and Gupta [2]. Going back to see what happened to the U.S. economy and environment in the period of time, we find that the highlight event, Saudi Arabia announced that it would increase its shareholding in Aramco to 60 percent in June 4th, 1974, may affect the U.S. stock markets.

4.2. Application to the air traffic data

We adopt the air traffic data in Hsu [6] to show the implementation of JEL based testing procedure in this section. Let T_i be the arrival time, $i = 1, \ldots, 213$. According to Hsu [6], the interarrival times $X_i = T_{i+1} - T_i$, $i = 1, \ldots, 212$ are independently exponential distributed. Our computational result shows that there is no change in the mean and variance of the interarrival times which matches the conclusions in Hsu [6] and Chen and Gupta [2].

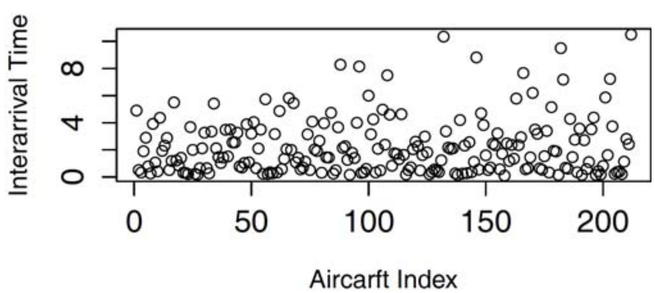

Figure 2. Interarrival times (in minutes) for the period from noon to 8:00 pm on April 30, 1969.

5. Discussion

In this article, a jackknife empirical likelihood based detection procedure for change-points in mean and variance is proposed. Simulations under different distribution settings with various sample sizes indicate that the power performance of the testing procedure is satisfactory. The proposed testing procedure is applied to the stock market data and the air traffic data. Significant change-points in the stock market data set are detected and there is no change detected in the air traffic data set. One may notice that the empirical likelihood estimators may not exist when the k or $n - k$ is too small. Therefore, the test statistic we proposed in Subsection 4.2 is a trimmed statistic which means that the testing procedure may not detect the change-point when it occurs in the very beginning or at the very end of data. Adjustment of the jackknife empirical likelihood based test is needed so that the testing procedure can detect all possible change locations. It is one of our ongoing research. Another ongoing work is to find the asymptotic distribution of the test statistic. A prospective candidate distribution is the extreme value distribution.

References

[1] J. Chen and A. K. Gupta, Change point analysis of a Gaussian model, Statist. Papers 40 (1999), 323-333.

[2] J. Chen and A. K. Gupta, Parametric Statistical Change Point Analysis: with Applications to Genetics, Medicine, and Finance, Birkhäuser, 2012.

[3] M. Csörgő and L. Horváth, Limit Theorems in Change-point Analysis, Wiley, New York, 1997.

[4] D. M. Hawkins, Testing a sequence of observations for a shift in location, J. Amer. Statist. Assoc. 72 (1977), 180-186.

[5] D. A. Hsu, Tests for variance shifts at an unknown time point, Appl. Stat. 26(3) (1977), 279-284.

[6] D. A. Hsu, Detecting shifts of parameter in gamma sequences with applications to stock price and air traffic flow analysis, J. Amer. Statist. Assoc. 74 (1979), 31-40.

[7] B. Y. Jing, J. Q. Yuan and W. Zhou, Jackknife empirical likelihood, J. Amer. Statist. Assoc. 104 (2009), 1224-1232.

[8] A. B. Owen, Empirical likelihood ratio confidence intervals for a single functional, Biometrika 75 (1988), 237-249.

[9] A. B. Owen, Empirical Likelihood, Chapman & Hall/CRC, New York, 2001.

[10] E. S. Page, Continuous inspection schemes, Biometrika 41 (1954), 100-116.

[11] E. S. Page, A test for a change in a parameter occurring at an unknown point, Biometrika 42 (1955), 523-547.

[12] P. Perron and T. J. Vogelsane, Testing for a unit root in a time series with a change mean - corrections and extensions, J. Bus. Econom. Statist. 10 (1992), 467-470.

[13] X. Shi, The approximate independence of jackknife pseudo-values and the bootstrap methods, Journal of Wuhan Institute Hydra-Electric Engineering 2 (1984), 83-90.

[14] L. J. Vostrikova, Detecting “disorder” in multidimensional random processes, Soviet Mathematics Doklady 24 (1981), 55-59.

[15] C. Zou, Y. Liu, P. Qin and Z. Wang, Empirical likelihood ratio test for the change-point problem, Statist. Probab. Lett. 77 (2007), 374-382.