

## Comments on: Extensions of some classical methods in change point analysis

Nirian Martín · Leandro Pardo

Published online: 9 May 2014  
© Sociedad de Estadística e Investigación Operativa 2014

Professor Horváth co-authored in 1997 the book entitled “Limit theorems in change-point analysis”. It can be seen as a milestone in the history of this topic that has a history of over 50 years. This review paper is very interesting, timely, highly informative and provides a comprehensive summary of the recent research on change-point analysis. This research has deserved considerable attention during the last years and is expected to develop new results in the near future. Professors Horváth and Rice are to be congratulated for this succinct and insightful presentation of the state-of-art of the change-point methodology. The review not only considers the most frequently used classical methods and ideas in change-point analysis but also some new methodology developed in the last years in relation to time series, multivariate, panel data and functional data. The applications of several classical results to real data sets permit us a better understanding as well as showing the importance and scope of this theory that was originated in the 1940s but is currently an active area of research. It is an almost an impossible task to review the very extensive literature on change-point analysis in a single paper. Nevertheless, the paper contains a very rich collection of the contributions not only from the classical results but also from the most recent papers in change-point analysis. We want to thank professors Horváth and Rice for their enormous effort in bringing us all this existing material on a topic of the highest relevance.

---

This comment refers to the invited paper available at doi:[10.1007/s11749-014-0368-4](https://doi.org/10.1007/s11749-014-0368-4).

---

N. Martín  
Department of Statistic, Carlos III University of Madrid, C/ Madrid 126, 28903 Getafe, Spain  
e-mail: [nirian.martin@uc3m.es](mailto:nirian.martin@uc3m.es)

L. Pardo (✉)  
Department of Statistic and O.R., Complutense University of Madrid,  
Plaza de Ciencias s/n, 28040 Madrid, Spain  
e-mail: [lpardo@mat.ucm.es](mailto:lpardo@mat.ucm.es)

The authors speak about several methods proposed in [Batsidis et al. \(2013\)](#) to detect changes in multinomial data based on divergence test statistics but the domain of application of divergence test statistics to detect changes goes far beyond that of multinomial data because divergence test statistics can be used for testing parameter changes in general models. In i.i.d. samples, the parametric approach based on likelihood was taken by many authors, see [Csörgö and Horváth \(1997\)](#) and [Chen and Gupta \(2000\)](#). So, we will discuss briefly the importance of divergence measures in detecting parameter changes in general populations.

### 1 Divergence-based procedure for parametric change-point analysis

Let  $X_1, X_2, \dots, X_K$  be  $K$  independent  $d$ -variate observations ( $d \in \mathbb{N}$ ) with distribution functions  $F_{\theta_i}(\mathbf{x}), \theta_i \in \mathbb{R}^m, i = 1, \dots, K, \mathbf{x} \in \mathbb{R}^d$ . In the following, we shall denote by  $f_{\theta}$  the probability density function associated to  $F_{\theta}$ . We are interested in testing  $H_0 : \theta_1 = \dots = \theta_K = \theta_0$  against  $H_1 : \theta_1 = \dots = \theta_{\kappa} = \theta_0 \neq \theta_{\kappa+1} = \dots = \theta_K$ . In the above formulation, apart from the parameters  $\theta_0$  and  $\theta_*$ , the location of the single change-point,  $\kappa \in \{1, \dots, K - 1\}$ , is unknown. [Hawkins \(1987\)](#) studied a weighted maximized Wald-type test statistic based on the difference of the maximum likelihood estimator (MLE) calculated from  $X_1, X_2, \dots, X_k$  and  $X_{k+1}, \dots, X_K$  for solving such a problem. The corresponding asymptotic null distribution was obtained via weak convergence of an empirical process to a standardized tied-down Bessel process, (2). Based on this result, [Batsidis et al. \(2014\)](#) studied the previous problem on the basis of phi-divergence test statistics. Let  $\widehat{\theta}_{0,k}^{(K)}$  be the MLE of  $\theta_0$ , based on the random sample  $X_1, \dots, X_k$  from distribution  $f_{\theta_0}$ , and let  $\widehat{\theta}_{*,k}^{(K)}$  be the MLE of  $\theta_*$ , based on the random sample  $X_{k+1}, \dots, X_K$  from distribution  $f_{\theta_*}$ . If hypothesis  $H_1$  is true, the probabilistic models  $f_{\widehat{\theta}_{0,k}^{(K)}}$  and  $f_{\widehat{\theta}_{*,k}^{(K)}}$  differ and the distance between  $f_{\widehat{\theta}_{0,k}^{(K)}}$  and  $f_{\widehat{\theta}_{*,k}^{(K)}}$  is large enough for some  $k \in \{1, \dots, K - 1\}$ . To establish statistical distances between  $f_{\widehat{\theta}_{0,k}^{(K)}}$  and  $f_{\widehat{\theta}_{*,k}^{(K)}}$ , we shall consider the  $\phi$ -divergence measure between  $f_{\widehat{\theta}_{0,k}^{(K)}}$  and  $f_{\widehat{\theta}_{*,k}^{(K)}}$  given by

$$D_{\phi}(f_{\widehat{\theta}_{0,k}^{(K)}}, f_{\widehat{\theta}_{*,k}^{(K)}}) = \int_{\mathcal{X}^{(d)}} f_{\widehat{\theta}_{*,k}^{(K)}}(\mathbf{x}) \phi \left( \frac{f_{\widehat{\theta}_{0,k}^{(K)}}(\mathbf{x})}{f_{\widehat{\theta}_{*,k}^{(K)}}(\mathbf{x})} \right) d\mathbf{x}$$

which can maximized in a weighted setting as follows

$$\epsilon T_{\phi}^{(K)} = \max_{k \in N^{(K)}(\epsilon)} T_{\phi}^{(K)}(k), \quad T_{\phi}^{(K)}(k) = \frac{k}{K} \left( 1 - \frac{k}{K} \right) D_{\phi}(f_{\widehat{\theta}_{0,k}^{(K)}}, f_{\widehat{\theta}_{*,k}^{(K)}}) \quad (1)$$

where  $N^{(K)}(\epsilon)$  is a subset of  $\{1, \dots, K - 1\}$ , conveniently chosen for  $\epsilon > 0$  small enough to obtain  $\frac{k}{K} \in [\epsilon, 1 - \epsilon]$ . The estimator of  $\kappa$  is given by  $\widehat{\kappa}_{\phi}^{(K)} = \arg \max_{k \in \{1, \dots, K\}} T_{\phi}^{(K)}(k)$ . Let  $\{W_{0,h}(t)\}_{t \in [0,1], h = 1, \dots, m}$ , denote  $m$  independent Brownian bridges. The asymptotic distribution of  $\epsilon T_{\phi}^{(K)}$  under the null hypothesis, as  $K$  goes to infinity, is given by

$$\sup_{t \in [\epsilon, 1-\epsilon]} \frac{1}{t(1-t)} \sum_{h=1}^m W_{0,h}^2(t), \quad (2)$$

and is valid to calibrate when the value of  $T_{\phi}^{(K)}(k)$  is large enough, for some  $k \in \{1, \dots, K-1\}$ , to reject to null hypothesis. In the simulation study of aforementioned paper, the test statistics based on the Hellinger distance were proposed as a good alternative with respect to different versions of the likelihood ratio test statistic when considering the closeness of the exact sizes to the nominal levels as well as the powers and this result is very appealing taking into account that the Hellinger distance has often been considered to be a tool for obtaining robust statistical techniques. For more details about  $\phi$ -divergence measures, see [Pardo \(2006\)](#) or [Cressie and Pardo \(2002\)](#).

It is well known that outliers affect test procedures leading to false conclusions. An interesting survey on robust change-point analysis can be seen in [Hušková \(2013\)](#). Divergence measures have been used not only to get divergence test statistics for detecting changes but also to get robust change-point estimators for different problems. [Lee and Park \(2001\)](#) considered a robust CUSUM test for the variance change in linear processes based on a trimming method, and demonstrated that it is necessary to use a robust method to prevent outliers from damaging the test procedure. Motivated by the viewpoint that the same phenomenon is anticipated to occur in other situations, [Lee and Na \(2005\)](#) considered a robust test for parameter changes based on the minimum density power divergence estimator (MDPDE). This estimator was introduced by [Basu et al. \(1998\)](#) for i.i.d. samples. Many density-based minimum distance procedures have been observed to have strong properties in estimation and testing together with high efficiency, e.g., [Basu et al. \(2011\)](#). The MDPDE has an advantage of not requiring smoothing methods and therefore one can avoid drawbacks and difficulties such as the selection of the bandwidth, that necessarily follows from the kernel smoothing method. [Lee and Na \(2005\)](#) extended the MDPDE to correlated observations. They construct a Wald-type test using the MDPDE and show that the test statistic converges weakly to the sup of a standard Brownian bridge under some regularity conditions. To establish that convergence, the strong convergence result of the estimator was proven. This Wald-type statistic constructed through estimators based on MDPDE not only constitutes a robust test against outliers, but also keeps the same efficiency as the MLE-based CUSUM test when there are no outliers. Recently, in [Basu et al. \(2013\)](#), new test statistics were proposed for some parametric hypothesis, based on density power divergences not only for estimation but also for constructing the test statistic. Since these test statistics exhibited strong robustness properties together with high efficiency, we think it might be interesting to explore them for change-point problems creating a link with the Wald-type test proposed by [Lee and Na \(2005\)](#).

## 2 Empirical divergence test statistics for change point analysis

Recently, new nonparametric methods based on empirical likelihood and likelihood ratio test have been proposed in the literature for detecting and estimating a change-

point from a sequence of random variables, see for instance Zou et al. (2007), Shen (2013) and references therein. Based on the results of Batsidis et al. (2013, 2014), it would be possible to consider general families (in the sense that these families would have as a particular case the empirical likelihood ratio test) of empirical divergence test statistics to solve different problems related to the change-point analysis.

**Acknowledgments** We would like to thank the Editors of TEST for their kind invitation to write this Comment. This research is partially supported by Grant MTM 2012-33740 from Ministerio de Economía y Competitividad (Spain).

## References

- Basu A, Harris IR, Hjort NL, Jones MC (1998) Robust and efficient estimation by minimizing a density power divergence. *Biometrika* 85:549–559
- Basu A, Shioya H, Park C (2011) *Statistical inference: the minimum distance approach*. Chapman & Hall/CRC, Boca Raton
- Basu A, Mandal A, Martín N, Pardo L (2013) Testing statistical hypotheses based on the density power divergence. *Ann Inst Stat Math* 65:319–348
- Batsidis A, Horváth L, Martín N, Pardo L, Zografos K (2013) Change-point detection in multinomial data using phi-divergence test statistics. *J Multivar Anal* 118:53–66
- Batsidis A, Martín N, Pardo L, Zografos K (2014) Phi-divergence based procedure for parametric change-point problems. *Methodol Comput Appl Probabl*. doi:[10.1007/s11009-014-9398-3](https://doi.org/10.1007/s11009-014-9398-3)
- Chen J, Gupta AK (2000) *Parametric statistical change point analysis*. Birkhäuser, Boston
- Cressie N, Pardo L (2002) Phi-divergence statistics. In: ElShaarawi AH, Pledger WW (eds.) *Encyclopedia of environmetrics*, vol 3. John Wiley & Sons, New York, pp 1551–1555
- Csörgö M, Horváth L (1997) *Limit theorems in change-point analysis*. Wiley, New York
- Hawkins DL (1987) A test for a change point in a parametric model based on a maximal wald-type statistic. *Sankhya A* 49:368–376
- Huškova M (2013) Robust change point analysis. In: *Robustness and complex data structures*. Springer, Berlin, pp 170–190
- Lee S, Park S (2001) The cusum of squares test for scale changes in infinite order moving average processes. *Scand J Stat* 30:781–796
- Lee S, Na O (2005) Test for parameter change based on the estimator minimizing density-based divergence measures. *Ann Inst Stat Math* 57:553–573
- Pardo L (2006) *Statistical inference based on divergence measures*. Chapman & Hall/CRC, Boca Raton
- Shen G (2013) On empirical likelihood inference of a change-point. *Stat Probabl Lett* 83:1662–1668
- Zou C, Liu Y, Qin P, Wang Z (2007) Empirical likelihood ratio test for the change-point problem. *Stat Probabl Lett* 77:374–382