Rejoinder on: Extensions of some classical methods in change point analysis

Lajos Horváth · Gregory Rice

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

The task of surveying a field as rich as change point analysis is difficult, if not impossible, to accomplish with a single paper. Therefore, to begin this rejoinder, we would like to thank the discussants for their insightful and valuable contributions in the form of comments on our paper. Many aspects of change point analysis which did not appear in our paper, or received little attention, have been given consideration and further references in the comments. Moreover, the points raised in the discussions illustrate the cutting edge of modern change point analysis and avenues for future research. Below we respond to the comments of each discussant individually and suggest related open problems.

Professors Hušková and Prašková focused their discussion on robust tests for change point detection, and they provided several references where robust estimators are used. So far univariate observations and parameters have comprised the bulk of the literature in this direction. Based on the comments of Professors Hušková and Prašková, it would be interesting and practical to consider the multivariate and infinite dimensional cases. This is a difficult question from a theoretical point of view. Presently, to our knowledge, not a single paper has been devoted to the effect of outliers in functional data. The estimation of the covariance function and the long run covariance function are average

Research supported by NSF grant DMS 1305858.

This rejoinder refers to the comments available at doi:[10.1007/s11749-014-0367-5](https://doi.org/10.1007/s11749-014-0367-5), doi:[10.1007/s11749-014-0369-3](https://doi.org/10.1007/s11749-014-0369-3), doi:[10.1007/s11749-014-0370-x](https://doi.org/10.1007/s11749-014-0370-x), doi:[10.1007/s11749-014-0371-9](https://doi.org/10.1007/s11749-014-0371-9), doi:[10.1007/s11749-014-0372-8](https://doi.org/10.1007/s11749-014-0372-8), doi:[10.1007/s11749-014-0373-7](https://doi.org/10.1007/s11749-014-0373-7), doi:[10.1007/s11749-014-0376-4](https://doi.org/10.1007/s11749-014-0376-4), doi:[10.1007/s11749-014-0377-3](https://doi.org/10.1007/s11749-014-0377-3).

L. Horváth (✉) · G. Rice
Department of Mathematics, University of Utah, Salt Lake City, UT 84112–0090, USA
e-mail: horvath@math.utah.edu

G. Rice
e-mail: rice@math.utah.edu

based, and hence the corresponding estimated eigenvalues and eigenfunctions are not robust. An interesting problem would be to find alternative robust estimators for these functions. The papers [Berkes and Horváth \(2012\)](#), [Berkes et al. \(2011\)](#) and [Berkes et al. \(2012\)](#) suggested the removal of the largest (by length) observations from the sample of univariate observations in an attempt to remove the outliers. The other interesting feature of this procedure is that CUSUM type processes based on the trimmed data will have limits derived from Gaussian processes even if the underlying errors have long tails. It is shown in [Berkes et al. \(2010\)](#) that a CUSUM process cannot be successfully resampled (bootstrapped or permutation sampled), but resampling works well for trimmed observations. These results assume independent and identically distributed errors and only [Bazarova et al. \(2012\)](#) considers a dependent case.

Also with regards to robust change point analysis, Professor Dehling in his discussion focused on the asymptotics of U and rank statistics under weak and strong dependence. For these statistics the application of weight functions often increases the power. The limit distribution is quite complicated in the strong dependence case and so resampling is suggested to get critical values. Of particular importance in economical and financial applications is to distinguish between change point models and long range dependence or heavy tails. One might, for example, conclude incorrectly based on the result of a change point test that a sequence exhibits strong dependence when instead there are changes in the mean. In such situations, it is possible that the reverse conclusion is reached by accounting for the change(s).

Professor Trapani makes a convincing argument for the importance of change point analysis in economics. Our survey could only provide a glimpse of further, more sophisticated models and possible applications. We assumed, for example, that in the panel data models the loading factor remained constant during the observation period. This is an assumption which should be tested. Also, one might want to replace μ_i with a linear regression. The extension of the methods presented in the paper to multivariate panels is straightforward. In financial applications high frequency data are common, and so functional panels would find applications and provide a framework for interesting theoretical developments.

Professor Kokoszka suggested the spatial model

$$X_n(\mathbf{s}_i, t) = \mu_n(\mathbf{s}_i, t) + \varepsilon_n(\mathbf{s}_i, t), \quad 1 \leq i \leq K, 0 \leq t \leq T, 1 \leq n \leq N,$$

$\varepsilon_n(\mathbf{s}, t) = 0$. In this model, we have functional observations at some locations $\mathbf{s}_i$, $1 \leq i \leq K$ and the sample size is N . Other possible test statistics could be based on the functionals of

$$M_k = \sum_{i=1}^K \int_0^1 \left(\sum_{\ell=1}^k X_{\ell}(\mathbf{s}_i, t) - \frac{k}{N} \sum_{\ell=1}^N X_{\ell}(\mathbf{s}_i, t) \right)^2 dt$$

which would provide a completely functional approach. An interesting problem is to obtain the asymptotic properties of functionals of M_k when N and K tend to infinity.

Professor Aston in his discussion brought up a fascinating aspect of change point analysis in higher dimensions; namely the issue of whether or not a change present in the data will be detected after a dimension reduction technique is applied.

A popular choice to perform such dimension reduction is to use the principle components generated by the long run covariance function of the data. Typically, there are several methods available to the practitioner to estimate the unknown covariance function, each of which leading to a consistent test after dimension reduction under the null hypothesis. Interestingly though, often times if the long run covariance is estimated in such a way that it is an inconsistent estimator under the alternative then an asymptotic power of one can be achieved. This cannot be guaranteed in general, as discussed by [Aston and Kirch \(2012\)](#), since if the data are projected into a space which is orthogonal to the change itself then the CUSUM test will not have power. These observations also are applicable in other situations, like testing the equality of the means of functional observations (cf. [Horváth and Rice 2013](#)). These issues can be avoided, at least for the simple change point problem, by employing a fully functional test, i.e., one which does not employ dimension reduction. An example of such a test is discussed in [Horváth et al. \(2014\)](#).

Professors Martín and Pardo provided an interesting outline of procedures based on divergence measures. The suggested methods are maximally selected two sample tests where the underlying distributions are in a parametric form; hence their method compares estimators when they are computed from the first k and last $N - k$ observations. An interesting new approach would be to use the divergence measures concept in a nonparametric setup, i.e., without assuming a parametric form for the underlying distributions. This would provide an alternative to empirical process-based procedures.

There were several comments on the possibility of multiple changes under the alternative, and this was the focus of the discussion written by Professor Berkes. Typically procedures derived under the exactly one change alternative have non-trivial power when there are multiple changes. If the number of changes and the structure of changes are known, more powerful tests are possible. The estimation of the number of changes has been widely investigated. Some interesting developments can be found in [Aue and Lee \(2011\)](#), [Kühn \(2001\)](#) and [Serbinowska \(1996\)](#), and such methods are computationally intensive. From a theoretical point of view, one needs an upper bound for the number of changes and a lower bound for the distance between consecutive change points. The approximation results discussed by Professor Berkes could be used to extend the results of [Kühn \(2001\)](#) to more complex models.

Online monitoring is crucial in several applications and, therefore, we agree with Professor Kirch that this is an important field of current and future research. Resampling could be useful to find “optimal” boundary functions for sequential change point procedures. The challenge in implementing such procedures is that the bootstrapping procedure must run quickly to be useful in many online monitoring situations. More empirical and theoretical research is needed in the field of online monitoring with dependent and/or high frequency data.

References

- Aston J, Kirch C (2012) Detecting and estimating changes in dependent functional data. *J Multivar Anal* 109(2012):204–220
- Aue A, Lee TCM (2011) On image segmentation using information theoretic criteria. *Ann Stat* 39:2912–2935

- Bazarova A, Berkes I, Horváth L (2012) Trimmed stable AR(1) processes. Preprint
- Berkes I, Horváth L (2012) The central limit theorem for sums of trimmed variables with heavy tails. *Stoch Process Appl* 122:449–465
- Berkes I, Horváth L, Schauer J (2010) Non-central limit theorems for random selections. *Probab Theory Relat Fields* 147:449–479
- Berkes I, Horváth L, Schauer J (2008) Permutation and bootstrap statistics under infinite variance. In: Doukhan P, Lang G, Surgailis D, Teyssire G (eds) *Dependence in probability and statistics proceedings of 2008 STATDEP. Lecture Notes in Statistics*, Springer, Berlin, pp 1–19
- Berkes I, Horváth L, Schauer J (2011) Asymptotics of trimmed CUSUM statistics. *Bernoulli* 17:1344–1367
- Berkes I, Horváth L, Schauer J (2012) Asymptotic behavior of trimmed sums. *Stoch Dyn* 12:1–14
- Horvth L, Kokoszka P, Rice G (2014) Testing stationarity of functional data. *J Econom* 179:66–82
- Horvth L, Rice G (2013) Testing equality of means when the observations are from functional time series. Preprint
- Khn C (2001) An estimator of the number of change points based on a weak invariance principle. *Stat Probab Lett* 51:189–196
- Serbinowska M (1996) Consistency of an estimator of the number of changes in binomial observations. *Stat Probab Lett* 29:337–344