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

Herold Dehling

First of all, I would like to compliment the authors for writing this very fine survey paper on some recent developments in change point analysis. There is a lot of current research activity devoted to change point analysis, and many articles have appeared since the publication of the seminal book by Csörgő and Horváth (1997). Among the topics covered in the present paper are empirical process techniques, Darling-Erdős laws, changes in correlations, changes in regression parameters, sequential testing, panel models, and functional data. Horváth and Rice provide an excellent survey of some of the recent results, which will be helpful to all researchers in this area.

In my discussion, I will complement the paper by Horváth and Rice by presenting some recent results on U-statistics based robust change-point tests for time series. In a series of papers, see Dehling and Fried (2012), Dehling, Fried, Garcia and Wendler (2013), Dehling, Rooch and Taqqu (2013a, 2013b), and Betken (2014), we have investigated such tests, and derived their asymptotic distribution, both in the case of short range as well as long range dependent time series.

We consider a model where the data are generated by $X_i = \mu_i + \epsilon_i$, where μ_i is an unknown signal, and where ϵ_i is a stationary ergodic noise process with $E(\epsilon_i) = 0$. Given the observations $X_1, \ldots, X_n$, we wish to test the hypothesis of no change, i.e. $H : \mu_1 = \ldots = \mu_n$, against the alternative

$$A : \mu_1 = \ldots = \mu_k \neq \mu_{k+1} = \ldots = \mu_n, \text{ for some } 1 \leq k \leq n-1.$$

H. Dehling
Fakultät für Mathematik
Ruhr-Universität Bochum
Universitätsstraße 150
44780 Bochum, Germany
E-mail: herold.dehling@rub.de

The motivation for our change-point tests, as for many other change-point tests, arises from the two-sample problem that one obtains when the change point k is assumed to be known. In this situation, we are given two samples $X_1, \ldots, X_k$, and $X_{k+1}, \ldots, X_n$, and we want to test for a difference in location. In this paper, we focus on three tests, namely the Gauss test, the Wilcoxon test and the Hodges-Lehmann test, which are associated with the test statistics

$$T_{n,1}(k) = \frac{1}{n-k} \sum_{i=k+1}^{n} X_i - \frac{1}{k} \sum_{i=1}^{k} X_i$$

$$T_{n,2}(k) = \sum_{i=1}^{k} \sum_{j=k+1}^{n} 1_{\{X_i \leq X_j\}}$$

$$T_{n,3}(k) = \text{median}\{(X_j - X_i) : 1 \leq i \leq k, k+1 \leq j \leq n\},$$

respectively. When the change point is unknown, we consider some summary statistic, e.g. a weighted maximum $\max_{1 \leq k \leq n-1} |a_n(k) T_{n,i}(k)|$.

In order to derive the asymptotic distribution of these change-point test statistics, we study weak convergence of the processes $(T_{n,i}([n\lambda]))_{0 \leq \lambda \leq 1}$, properly normalized. For the first two statistics, this leads to the two-sample U-statistic process

$$U_{[n\lambda], n-[n\lambda]} = \frac{1}{[n\lambda](n-[n\lambda])} \sum_{i=1}^{[n\lambda]} \sum_{j=[n\lambda]+1}^{n} h(X_i, X_j), \ 0 \leq \lambda \leq 1.$$

The Hodges-Lehmann estimator is the median of the empirical distribution of the pairwise differences $X_j - X_i, 1 \leq i \leq k < j \leq n$. Thus, we are lead to the study of the empirical distribution and the empirical quantiles of arbitrary kernels $g(X_i, X_j), 1 \leq i \leq k < j \leq n$. We define the empirical distribution function

$$U_{[n\lambda], n-[n\lambda]}(t) = \frac{1}{[n\lambda](n-[n\lambda])} \sum_{i=1}^{[n\lambda]} \sum_{j=[n\lambda]+1}^{n} 1_{\{g(X_i, X_j) \leq t\}}, \ 0 \leq \lambda \leq 1, t \in \mathbb{R},$$

and the quantile function $Q_{[n\lambda], n-[n\lambda]}(p) = U_{[n\lambda], n-[n\lambda]}^{-1}(p)$, where U^{-1} denotes the generalized inverse.

We have obtained results both under the assumption of Short Range Dependent (SRD) as well as of Long Range Dependent (LRD) noise. We will first focus on the SRD case, where we assume that the noise has a representation as a functional of a β-mixing process $(Z_i)_{i \in \mathbb{Z}}$, i.e. that $\epsilon_i = f(Z_i, Z_{i-1}, \ldots)$, where f is a Lipschitz continuous function. This covers all relevant examples from time series analysis, such as ARMA and GARCH models. In addition, many deterministic dynamical systems such as expanding maps of the unit interval are covered as well.

Dehling, Fried, Garcia and Wendler (2013) have studied the asymptotic distribution of the two-sample U-statistics process in the case of SRD noise,

extending results obtained earlier by Csörgő and Horváth (1988) in the case of IID noise. Under some technical conditions, concerning the β-mixing coefficients and the continuity of f and h, and under the null hypothesis of no change, we could show that

$$\sqrt{n}\lambda(1-\lambda)(U_{[n\lambda],n-[n\lambda]}-\theta)_{0\leq\lambda\leq 1} \xrightarrow{\mathcal{D}} ((1-\lambda)W_1(\lambda)+\lambda(W_2(1)-W_2(\lambda)))_{0\leq\lambda\leq 1},$$

where (W_1, W_2) is 2-dimensional Brownian motion with covariance structure

$$\text{Cov}(W_i(\lambda), W_j(\lambda)) = (\lambda \wedge \mu) \sum_{k\in\mathbb{Z}} \text{Cov}(h_i(X_0), h_j(X_k)),$$

for $i, j \in \{1, 2\}$ and $0 \leq \lambda, \mu \leq 1$. The functions h_i are the first order terms in the Hoeffding decomposition of the kernel h, given by $h(x, y) = \theta + h_1(x) + h_2(y) + \psi(x, y)$. Here, $\theta = Eh(X, Y)$, $h_1(x) = Eh(x, Y) - \theta$, $h_2(y) = Eh(X, y) - \theta$ and $\psi(x, y) = h(x, y) - \theta - h_1(x) - h_2(y)$, where X, Y are independent random variables with the same distribution as X_1.

Applying the above limit theorem for the two-sample U-process to the kernels $h(x, y; t) = 1_{\{g(x,y)\leq t\}}$, $t \in \mathbb{R}$, one can obtain convergence of the two-sample empirical U-process $(\sqrt{n}\lambda(1 - \lambda)(U_{[n\lambda],n-[n\lambda]}(t) - U(t)))_{0\leq\lambda\leq 1}$, where $U(t) = P(g(X, Y) \leq t)$, for fixed t. Using a Bahadur-Kiefer representation, one can then obtain convergence of the corresponding two-sample quantile process. In a forthcoming paper, Dehling, Fried and Wendler (2014) show that, again under some technical conditions, concerning the β-mixing coefficients and the continuity of f, g and U, and under the null hypothesis of no change, the two-sample quantile process

$$\sqrt{n}\lambda(1 - \lambda) \left(Q_{[n\lambda],n-[n\lambda]}(p) - Q(p) \right)_{0\leq\lambda\leq 1}$$

converges in distribution to the process $((1-\lambda)W_1(\lambda)+\lambda(W_2(1)-W_2(\lambda)))_{0\leq\lambda\leq 1}$, where $(W_1(\lambda), W_2(\lambda))$ is 2-dimensional Brownian motion with covariance function

$$\text{Cov}(W_i(\mu), W_j(\lambda)) = \frac{\mu \wedge \lambda}{u^2(Q(p))} \sum_{k\in\mathbb{Z}} \text{Cov}(h_i(X_0, Q(p)), h_j(X_k, Q(p))).$$

Here, $Q(p) = U^{-1}(p)$ denotes the quantile function. Moreover, we denote the terms of the Hoeffding decomposition of $h(x, y; t)$ by $h_1(x; t)$ and $h_2(y; t)$, and define $u(t) = U'(t)$.

When applying this result to the case of the Hodges-Lehmann change point test, we can show that

$$\sqrt{n} \max_{1\leq k\leq n} \frac{k}{n}(1 - \frac{k}{n})| \text{median}\{(X_j - X_i) : 1 \leq i \leq k < j \leq n\}|$$

has asymptotically the same distribution as $\frac{\sigma}{u(0)} \sup_{0\leq\lambda\leq 1} |W^{(0)}(\lambda)|$, where $\sigma^2 = \sum_{k=-\infty}^{\infty} \text{Cov}(F(X_0), F(X_k))$, and where $(W^{(0)}(\lambda))_{0\leq\lambda\leq 1}$ denotes standard Brownian bridge.

Dehling, Rooch and Taqqu (2013a) have studied the asymptotic distribution of the Wilcoxon change-point test in the case of LRD noise. We specifically consider Gaussian subordinated processes, i.e. we assume that $\epsilon_i = H(\xi_i)$, where $(\xi_i)_{i \geq 1}$ is a stationary Gaussian process with standard normal marginals and autocorrelation function $\rho_k = k^{-D}L(k)$, $0 < D < 1$, L slowly varying, and where H is some measurable real-valued function. Moreover, we define $J_k(x) = E(1_{\{H(\xi) \leq x\}}H_k(\xi))$, where H_k is the k-th order Hermite polynomial. The smallest integer m such that $J_m(x) \not\equiv 0$ is called the Hermite rank. Define the normalizing constants $d_n = \mathrm{Var}(\sum_{i=1}^{n} H_m(\xi_i))$, and recall that $d_n \sim cn^{2H}L^m(n)$, where $H := 1 - mD/2$ is the Hurst coefficient.

Assuming that $m < 1/D$, Dehling, Rooch and Taqqu (2013) showed that, under the null hypothesis of no change, the process

$$\frac{1}{n\,d_n}\left(\sum_{i=1}^{[n\lambda]}\sum_{j=[n\lambda]+1}^{n}\left(1_{\{X_i \leq X_j\}} - \frac{1}{2}\right)\right)_{0 \leq \lambda \leq 1}$$

converges in distribution to the proces $\frac{\int J_m(x)dF(x)}{m!}(Z_m(\lambda) - \lambda Z_m(1))_{0 \leq \lambda \leq 1}$, where $(Z_m(\lambda))_{0 \leq \lambda \leq 1}$ denotes an m-th order Hermite process. In a recent paper, Betken (2014) gives a self-normalized version of the Wilcoxon change-point test which is asymptotically distribution free under the null hypothesis.

In a subsequent paper, Dehling, Rooch and Taqqu (2013b) have investigated the asymptotic distribution of the Wilcoxon and the CUSUM change point test under local alternatives, and calculated their asymptotic relative efficiencies. In the case of Gaussian errors, the ARE equals 1, while for heavy-tailed data, the Wilcoxon test has superior power. For finite samples, these results are confirmed by simulations.

References

1. A. Betken: Testing for change-points in long-range dependent time series by means of a self-normalized Wilcoxon test, *arXiv:1403.0265 [math.ST]* (2014).
2. M. Csörgő and L. Horváth: Invariance Principles for Changepoint Problems. *Journal of Multivariate Analysis* **27**, 151–168 (1988).
3. M. Csörgő and L. Horváth: *Limit Theorems in Change-Point Analysis*, Wiley, Chichester (1997).
4. H. Dehling and R. Fried: Asymptotic Distribution of Two-Sample Empirical U-Quantiles with Applications to Robust Tests for Shifts in Location, *Journal of Multivariate Analysis* **105**, 124–140 (2012).
5. H. Dehling, R. Fried, I. Garcia and M. Wendler: Change-Point Detection under Dependence Based on Two-Sample U- Statistics, To appear in: *Asymptotic Methods in Stochastics - Festschrift in Honor of Miklos Csörgő's 80th Birthday arXiv:1304.2479 [math.ST]*; (2013).
6. H. Dehling, A. Rooch and M. S. Taqqu: Nonparametric Change-Point Tests for Long-Range Dependent Data, *Scandinavian Journal of Statistics* **40**, 153–173 (2013)
7. H. Dehling, A. Rooch and M.S. Taqqu: Power of Change-Point Tests for Long-Range Dependent Data, *arXiv:1303.4917 [math.ST]* (2013).