Change-point Problems
IMS Lecture Notes – Monograph Series (Volume 23, 1994)

NONPARAMETRIC CHANGE-POINT TESTS OF THE KOLMOGOROV-SMIRNOV TYPE

BY DIETMAR FERGER
University of Giessen

We consider a triangular array $X_1^n, \cdots, X_n^n$, $n \in \mathbb{N}$, of row-wise independent random elements with values in a measurable space $(\mathcal{X}, \mathcal{B})$. Suppose there exists a $\theta \in [0,1]$ such that $X_1^n, \cdots, X_{[n\theta]}^n$ have distribution ν_1 and $X_{[n\theta]+1}^n, \cdots, X_n^n$ have distribution ν_2. We construct tests for $H_0^{(1)} : \theta \in \{0,1\}$ versus $H_1^{(1)} : \theta \in (0,1)$, $H_0^{(2)} : \theta = \theta_0$, $\theta_0 \in (0,1)$, versus $H_1^{(2)} : \theta \neq \theta_0$ and $H_0^{(3)} : \theta \in \Theta_0$ versus $H_1^{(3)} : \theta \notin \Theta_0$, where Θ_0 is a closed subset of $(0,1)$. The tests, which are based on U-statistics type processes, are shown to be asymptotic level-α tests and consistent on a large class of alternatives. For $H_0^{(1)}$ versus $H_1^{(1)}$ a careful investigation of the power function is provided. The results are part of the author's (1991) dissertation written under the supervision of Professor Stute. Proofs and more detailed information will be published elsewhere.

1. Introduction and Main Results. Let $X_1^n, \cdots, X_n^n$, $n \in \mathbb{N}_1$ be row-wise independent random elements defined on a probability space $(\Omega, \mathcal{A}, P)$ with values in a measurable space $(\mathcal{X}, \mathcal{B})$. Suppose there exists $\theta \in [0,1)$ such that $X_1^n, \cdots, X_{[n\theta]}^n$ have distribution ν_1 and $X_{[n\theta]+1}^n, \cdots, X_n^n$ have distribution $\nu_2 \neq \nu_1$, where both ν_1 and ν_2 as well as the *change-point* θ are unknown. In the *standard* test problem $H_0^{(1)} : \theta = 0$ versus $H_1^{(1)} : \theta \in (0,1)$ we ask whether there is a change at all, whereas in the *nonstandard* test situation, we want to know if a change has taken place at a certain point (or within a certain time interval) or not. Formally, we are interested in testing $H_0^{(2)} : \theta = \theta_0, \theta_0 \in (0,1)$, versus $H_1^{(2)} : \theta \neq \theta_0$ and $H_0^{(3)} : \theta \in \Theta_0$ versus $H_1^{(3)} : \theta \notin \Theta_0$, where $\Theta_0 \subseteq (0,1)$, respectively. All tests we recommend are based on a stochastic process first introduced by Csörgő and Horváth (1988) in the case of real-valued data:

$$r_n(t) = n^{-2} \sum_{i=[nt]+1}^{n} \sum_{j=1}^{[nt]} K(X_i^n, X_j^n), \qquad 0 \le t \le 1,$$

AMS 1991 Subject Classification: Primary 62G10, 62G20; Secondary 60G15, 62G09.

Key words and phrases: Tests for nonstandard change-point problems, consistency, power function, bootstrap approximation.

where $K : \mathcal{X}^2 \to \mathbb{R}$ is a measurable kernel. In what follows K is assumed to be bounded. This guarantees the existence of all integrals under no assumptions on the underlying distributions. For the motivation of our test the following result of Ferger and Stute (1992) turns out to be very useful:

$$\sup_{0\le t\le 1} |r_n(t) - r(t)| = O\left[\left\{\frac{\log n}{n}\right\}^{1/2}\right] \tag{1}$$

with probability one (w.p.1) $\forall\ \theta \in [0,1)$, where

$$r(t) \equiv r_{\mu,\tau,\lambda,\theta}(t) = \begin{cases} t\mu(\theta - t) + t\lambda(1-\theta), & 0 \le t \le \theta \\ (1-t)\lambda\theta + (1-t)\tau(t-\theta), & \theta < t \le 1 \end{cases}$$

with $\mu = \int K d\nu_1 \otimes \nu_1$, $\tau = \int K d\nu_2 \otimes \nu_2$ and $\lambda = \int K d\nu_2 \otimes \nu_1$. We will only consider *antisymmetric* kernels $K : K(x,y) = -K(y,x)\ \forall\ x, y \in \mathcal{X}$. Note that in this case $\mu = \tau = 0$ and therefore $r \equiv r_{\lambda,\theta}$ vanishes on the whole unit interval, if $\theta = 0$. Thus by (1) we have that with probability one

$$n^{1/2} \sup_{0\le t\le 1} |r_n(t)| \begin{cases} = O((\log n)^{1/2}), & \text{if } H_0^{(1)} \text{ holds} \\ \ge \text{const} \cdot n^{1/2}, & \text{if } H_1^{(1)} \text{ holds with } \lambda \ne 0. \end{cases} \tag{2}$$

So, large values of the left hand side in (2) indicate a change in the distribution. In the case $\mathcal{X} = \mathbb{R}$, Csörgő and Horváth (1988) proved that $(n^{1/2} r_n)_{n\in\mathbb{N}}$ converges in distribution to σB_0 under $H_0^{(1)}$, where B_0 denotes a Brownian Bridge and $\sigma^2 \equiv \sigma^2(K,\nu_2) = \int \left[\int K(x,y)\nu_2(dx)\right]^2 \nu_2(dy)$. Their invariance principle can be extended to arbitrary measurable spaces. We note that sometimes $\sigma^2 = \sigma^2(K,\nu_2)$ is known. For example, if $\mathcal{X} = \mathbb{R}$, $K(x,y) = \operatorname{sign}(x-y)$ and ν_2 is continuous, then $\sigma^2 = 1/3$. However, in general, σ^2 is not known. In this case it is possible to define an estimator σ_n^2 such that σ_n^2 converges to σ^2 w.p.1, if $H_0^{(1)}$ holds. Moreover, we prove that under $H_1^{(1)}$ $\sigma_n^2 \to \sigma^2(K,\nu_\theta)$ w.p.1, where $\nu_\theta = \theta\nu_1 + (1-\theta)\nu_2$, if in addition K is continuous and $\mathcal{X}$ is a separable metric space. Herewith we are in the position to generalize the Csörgő–Horváth (1988) test. Let α denote the error of the first kind, then the test is defined by

$$t_{n,\alpha}^{(1)} = 1_{\{\sigma_n^{-1} n^{1/2} \sup_{0\le t\le 1} |r_n(t)| > c_1(\alpha)\}},$$

where $c_1(\alpha)$ is the $(1-\alpha)$-quantile of the Kolmogorov-Smirnov distribution.

THEOREM 1. *Let $\mathcal{X}$ be a separable metric space and let K be a bounded, antisymmetric and continuous kernel. Then* $\lim_{n\to\infty} P_0(t_{n,\alpha}^{(1)} = 1) = \alpha\ \forall\ \theta \in (0,1]$ *and* $\lim_{n\to\infty} P_\theta(t_{n,\alpha}^{(1)} = 1) = 1\ \forall\ \alpha \in (0,1]\ \forall\ \theta \in (0,1)\ \forall\ \nu_1, \nu_2$ *with* $\lambda \ne 0$.

Next, we are interested in the *power* of our test. Let $\gamma_n^{(\alpha)}(\theta,\nu_1,\nu_2) := P_{\theta,\nu_1,\nu_2}(t_{n,\alpha}^{(1)}=1)$ be the *power function*.

THEOREM 2. *Under the same assumption as in Theorem 1,* $\lim_{n\to\infty}\gamma_n^{(\alpha)}(n^{-1/2},\nu_1,\nu_2)=\beta_\sigma(\lambda)$, *where* $\sigma^2=\sigma^2(K,\nu_2)$ *and* $\beta_\sigma(\lambda)$ *is an analytical completely known function. Furthermore, if* $(\nu_{2,n})_{n\in\mathbb{N}}$ *converges weakly to* ν_1 *in such a way that* $\lambda_n=\int K d\nu_{2,n}\otimes\nu_1 \sim n^{-1/2}$, *then* $\lim_{n\to\infty}\gamma_n^{(\alpha)}(\theta,\nu_1,\nu_{2,n})=1-h_{\sigma,\theta}(c_1(\alpha))\ \forall\ \theta\in[0,1)$, *where* $\sigma^2=\sigma^2(K,\nu_1)$ *and* $h_{\sigma,\theta}(c_1(\alpha))$ *is a certain boundary crossing probability of the Brownian Bridge.*

We have determined the boundary crossing probability of Theorem 2 explicitly. For $H_0^{(2)}:\theta=\theta_0$ versus $H_1^{(2)}:\theta\neq\theta_0$ with $0<\theta_0<1$ given, we proceed as in the standard case and propose the following test:

$$t_{n,\alpha}^{(2)} = 1_{\{n^{1/2}\sup_{0\le t\le 1}|r_n(t)-r_{\lambda_n,\theta_0}(t)|>c_2(\alpha)\}}.$$

Here, $\lambda_n=r_n(\theta_0)/(\theta_0(1-\theta_0))$ and $c_2(\alpha)$ is the $(1-\alpha)$-quantile of a centered Gaussian process, which may be specified. Since these quantiles are, if at all, not easily available, one has to replace them by a bootstrap approximation of the *exact* critical value, namely the $(1-\alpha)$-quantile of $n^{1/2}\sup_{0\le t\le 1}|r_n(t)-r_{\lambda_n,\theta_0}(t)|$. We can prove consistency of these quantiles.

THEOREM 3. *Let $\mathcal{X}$ be a measurable space and let K be a bounded and antisymmetric kernel. Then* $\lim_{n\to\infty}P_{\theta_0}(t_{n,\alpha}^{(2)}=1)=\alpha$ *and* $\lim_{n\to\infty}P_\theta(t_{n,\alpha}^{(2)}=1)=1\ \forall\ \theta\neq\theta_0\ \forall\ \nu_1,\nu_2$ *with* $\lambda\neq 0$.

In the more general situation of $H_0^{(3)}:\theta\in\Theta_0$ versus $H_1^{(3)}:\theta\notin\Theta_0$, the test $t_{n,\alpha}^{(2)}$ is not applicable, because it involves the quantities λ_n and θ_0, which are now unknown. Indeed, neither under $H_0^{(3)}$ nor under $H_1^{(3)}$ we know the value θ_0 of the change-point. However, following Ferger and Stute (1992) we can define an estimator θ_n such that $\theta_n-\theta_0=0(n^{-1}\log n)$ w.p.1 under $H_0^{(3)}$ and $\theta_n\to\theta_1\neq\theta_0$ w.p.1, if $H_1^{(3)}$ holds. Therefore we suggest:

$$t_{n,\alpha}^{(3)} = 1_{\{n^{1/2}\sup_{0\le t\le 1}|r_n(t)-r_{\hat\lambda_n,\theta_n}(t)|>c_2(\alpha)\}},$$

where $\hat\lambda_n=r_n(\theta_n)/(\theta_n(1-\theta_n))$.

THEOREM 4. *Assume θ_0 is closed and the assumptions of Theorem 3 are satisfied. Then* $\lim_{n\to\infty}P_\theta(t_{n,\alpha}^{(3)}=1)=\alpha\ \forall\ \theta\in\Theta_0$ *and* $\lim_{n\to\infty}P_\theta(t_{n,\alpha}^{(3)}=1)=1\ \forall\ \theta\notin\Theta_0\ \forall\ \nu_1,\nu_2$ *with* $\lambda\neq 0$.

REMARKS.

(1) We prefer to state our results for bounded kernels K, since in this case the conditions are completely carried by the given K rather than by the

unknown distributions ν_1 and ν_2. Moreover, tests induced by bounded kernels will be robust against outliers. Nevertheless, boundedness of K is not necessary and may be replaced by some moment conditions.

(2) Clearly, our approach raises the question, how to choose an *optimal* kernel.

(3) In the standard framework Csörgő and Horváth (1988) proposed to deal with *weighted* versions of $r_n(t)$, in order to obtain tests, which are more sensitive in the tails.

REFERENCES

CSÖRGŐ, M. and HORVÁTH, L. (1988). Invariance principles for change-point problems. *J. Multivariate Anal.* **27**, 151–168.

FERGER, D. (1991). Nonparametric change-point detection based on U-statistics. Dissertation, University of Giessen.

FERGER, D. and STUTE, W. (1992). Convergence of change-point estimators. *Stochastic Process. Appl.* **42**, 345–351.

MATHEMATICAL INSTITUTE
UNIVERSITY OF GIESSEN
ARNDSTRASSE 2
D-35392 GIESSEN, GERMANY