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

WEIGHTED MULTIVARIATE EMPIRICAL PROCESSES AND CONTIGUOUS CHANGE-POINT ANALYSIS

BY MIKLÓS CSÖRGŐ AND BARBARA SZYSZKOWICZ
Carleton University

Let $X = (X^{(1)}, \cdots, X^{(d)})$, $X_i = (X_i^{(1)}, \cdots, X_i^{(d)})$, $i = 1, 2, \cdots$, be independent random vectors in $\mathbb{R}^d$, $d \geq 1$. When testing for the possibility of having a change in the distribution of a sequence of *chronologically ordered* d-dimensional observations $X_i = (X_i^{(1)}, \cdots, X_i^{(d)})$, $i = 1, \cdots, n$, at an unknown time $1 \leq k < n$, it is natural to compare the empirical distributions "before" to those "after", via studying the asymptotic distribution of the sequence of statistics

$$\sup_{1 \leq k < n} \sup_{\mathbf{x} \in \mathbb{R}^d} n^{1/2} \left| \frac{1}{k} \sum_{i=1}^{k} \mathbf{1}(X_i \leq \mathbf{x}) - \frac{1}{n-k} \sum_{i=k+1}^{n} \mathbf{1}(X_i \leq \mathbf{x}) \right|$$

$$= \sup_{1 \leq k < n} \sup_{\mathbf{x} \in \mathbb{R}^d} \left| \sum_{i=1}^{k} \mathbf{1}(X_i \leq \mathbf{x}) - \frac{k}{n} \sum_{i=1}^{n} \mathbf{1}(X_i \leq \mathbf{x}) \right| \Big/ \left(n^{1/2} \left(\frac{k}{n} \left(1 - \frac{k}{n} \right) \right) \right).$$

These statistics however converge in distribution to ∞, as $n \to \infty$, even if the null assumption of no change in distribution were true. This is due to the weight function $((k/n)(1 - k/n))$ converging too fast to zero as $k/n \to 0$ and $k/n \to 1$. This remains true even if we were to replace this function by $((k/n)(1-k/n))^{1/2}$, $1 \leq k < n$. Thus we are led to considering multi-time parameter empirical processes with weights which would continue emphasizing the possibility of having a change in distribution, but in a non-degenerate way. Proofs of our results and further details will be given in a paper which is in preparation by the authors for publication elsewhere. This is an extended abstract of this forthcoming work.

1. Introduction. For an arbitrary, right continuously defined distribution function H on $\mathbb{R}^1$ we define the inverse (quantile function) of H by

$$H^{-1}(y) = \inf\{x \in \mathbb{R}^1 : H(x) \geq y\},\ 0 < y \leq 1,\ H^{-1}(0) = H^{-1}(0+).$$

Let $F_{(j)}(x_j)$,, $1 \leq j \leq d$, be the jth marginal of $F(\mathbf{x}) = F(x_1, \cdots, x_d)$, $\mathbf{x} = (x_1, \cdots, x_d) \in \mathbb{R}^d$, $d \geq 1$, and let $F_{(j)}^{-1}(u_j)$, $1 \leq j \leq d$, $\mathbf{u} = (u_1, \cdots, u_d) \in$

AMS 1991 Subject Classification: Primary 60F17, 62G30; Secondary 62G10, 60F25.

Key words and phrases: Multivariate empirical processes, multi-time parameter Gaussian processes, contiguity, change in distribution, testing for independence.

$I^d := [0,1]^d$, $d \geq 1$, be the inverses (quantile functions) of these marginals of F on $\mathbb{R}^d$.

Define the map $L^{-1} : I^d \to \mathbb{R}^d$ by

$$L^{-1}(\mathbf{u}) = L^{-1}(u_1, \cdots, u_d) := (F_{(1)}^{-1}(u_1), \cdots, F_{(d)}^{-1}(u_d)), \quad \mathbf{u} = (u_1, \cdots, u_d) \in I^d.$$

Then the map $L : \mathbb{R}^d \to I^d$, defined by

$$L(\mathbf{x}) = L(x_1, \cdots, x_d) := (F_{(1)}(x_1), \cdots, F_{(d)}(x_d)), \quad \mathbf{x} = (x_1, \cdots, x_d) \in \mathbb{R}^d,$$

is an *inverse* to L^{-1} in that, for $\mathbf{x} \in \mathbb{R}^d$ and $\mathbf{u} \in (0,1)^d$, we have the *component wise inequalities* $L(\mathbf{x}) \geq \mathbf{u}$ if and only if $\mathbf{x} \geq L^{-1}(\mathbf{u})$, and $L(L^{-1}(\mathbf{u})-) \leq \mathbf{u} \leq L(L^{-1}(\mathbf{u}))$. Consequently, if the components of L are continuous, then $L(L^{-1}(\mathbf{u})) = \mathbf{u}$.

We define the $(d+1)$-time parameter empirical process $\beta_n(\mathbf{x}, t)$ by

$$\beta_n(\mathbf{x}, t) = n^{-1/2} \sum_{i=1}^{nt} (\mathbf{1}(X_i \leq \mathbf{x}) - F(\mathbf{x})), \quad \mathbf{x} \in \mathbb{R}^d, \ 0 \leq t \leq 1.$$

A *Kiefer process* $\{K_F(\mathbf{x}, t), \mathbf{0} \leq L(\mathbf{x}) \leq \mathbf{1}, t \geq 0\}$ *on* $\mathbb{R}^d \times [0, \infty)$ *associated with a distribution function* F *on* $\mathbb{R}^d$, $d \geq 1$, is a separable $(d+1)$-parameter real valued Gaussian process with $K_F(\mathbf{x}, 0) = 0$, $EK_F(\mathbf{x}, t) = 0$ and

$$EK_F(\mathbf{x}, s)K_F(\mathbf{y}, t) = (s \wedge t)(F(\mathbf{x} \wedge \mathbf{y}) - F(\mathbf{x})F(\mathbf{y})).$$

From now on we assume that F on $\mathbb{R}^d$ is continuous and summarize some of our results in this case. Analogous results hold true when F is not assumed to be continuous. For similar studies when $d = 1$, we refer to Deshayes and Picard (1986) and Szyszkowicz (1991).

We will write

$$\{\alpha_n(\mathbf{u}, t), \ \mathbf{u} \in I^d, \ 0 \leq t \leq 1\} := \{\beta_n(L^{-1}(\mathbf{u}), t) \ \mathbf{u} \in I^d, \ 0 \leq t \leq 1\}.$$

By letting $G(\mathbf{u}) = F(L^{-1}(\mathbf{u}))$, with $\mathbf{x} = L^{-1}(\mathbf{u})$, $\mathbf{u} \in I^d$, $d \geq 1$, we have

$$\{K_F(\mathbf{x}, t), \ \mathbf{0} \leq L(\mathbf{x}) \leq \mathbf{1}, \ t \geq 0\} = \{K_G(\mathbf{u}, t), \ \mathbf{u} \in I^d, \ t \geq 0\}.$$

By THEOREM of Csörgő and Horváth (1988), associated with an F on $\mathbb{R}^d$, there is a Kiefer process $\{K_G(\mathbf{u}, t), \ \mathbf{u} \in I^d, \ t \geq 0\}$ such that, as $(nt) \to \infty$, $0 < t \leq 1$, we have

$$\sup_{\mathbf{u} \in I^d} |n^{1/2}\alpha_n(\mathbf{u}, t) - K_G(\mathbf{u}, nt)| \stackrel{\text{a.s}}{=} O((nt)^{1/2 - 1/(4d)}(\log(nt))^{3/2}).$$

With $U_i = L(X_i)$ then, in the continuous case we define the "tied down in t" multi-time parameter empirical bridge $\hat{\alpha}_n(\mathbf{s}, t)$ by

$$\hat{\alpha}_n(\mathbf{s},t) = \begin{cases} n^{-1/2}\left(\sum_{i=1}^{[(n+1)t]} \mathbf{1}(U_i \le \mathbf{s}) - \frac{[(n+1)t]}{n}\sum_{i=1}^{n}\mathbf{1}(U_i \le s)\right), & 0 \le t < 1, s \in I^d \\ 0, & t = 1, \mathbf{s} \in I^d, \end{cases}$$

and the Gaussian process $\{\Gamma_G(\mathbf{s}, t),\ \mathbf{s} \in I^d,\ 0 \le t < 1\}$ by

$$\Gamma_G(\mathbf{s},t) = K_G(\mathbf{s},t) - tK_G(\mathbf{s},1),\ \ \mathbf{s} \in I^d,\ \ 0 \le t < 1. \tag{1}$$

Consequently, $\Gamma_G(\cdot,\cdot)$ is a mean zero Gaussian process with covariance function

$$\begin{aligned} E\Gamma_G(\mathbf{s}_1,t_1)\Gamma_G(\mathbf{s}_2,t_2) &= \Gamma_G(\mathbf{s}_1 \wedge \mathbf{s}_2) - G(\mathbf{s}_1)G(\mathbf{s}_2))(t_1 \wedge t_2 - t_1t_2) \\ &= (F(L^{-1}(\mathbf{s}_1 \wedge \mathbf{s}_2)) - F(L^{-1}(\mathbf{s}_1))F(L^{-1}(\mathbf{s}_2)))(t_1 \wedge t_2 - t_1t_2). \end{aligned} \tag{2}$$

THEOREM 1. *Assume that $X_1, X_2, \cdots$ are independent with continuous distribution function F on $\mathbb{R}^d$. Then, there exists a Kiefer process $\{K_G(\mathbf{s}, t),\ \mathbf{s} \in I^d,\ t \ge 0\}$ such that with*

$$\begin{aligned} \{\Gamma_{G,n}(\mathbf{s},t), \mathbf{s} \in I^d,\ 0 \le t \le 1\} &:= \{n^{-1/2}(K_G(\mathbf{s},nt) - tK_G(\mathbf{s},n)), \\ &\qquad \mathbf{s} \in I^d,\ 0 \le t \le 1\} \\ &\overset{\mathcal{D}}{=} \{\Gamma_G(\mathbf{s},t),\ \mathbf{s} \in I^d,\ 0 \le t \le 1\} \text{ for each } n \ge 1, \end{aligned}$$

and a weight function q, which is positive on $(0,1)$ (i.e., $\inf_{\delta \le t \le 1-\delta} q(t) > 0$ for all $0 < \delta \le 1/2$), we have:

(a) *if*
$$\lim_{t \downarrow 0, t \uparrow 1} \left(t(1-t)\log\log\frac{1}{t(1-t)}\right)^{1/2} / q(t) = 0$$

then, as $n \to \infty$

$$\sup_{0<t<1}\sup_{\mathbf{s} \in I^d} \left|\hat{\alpha}_n(\mathbf{s},t) - \Gamma_{G,n}(\mathbf{s},t)\right| / q(t) = o_P(1), \tag{3}$$

(b) *if*
$$\limsup_{t \downarrow 0, t \uparrow 1} \left(t(1-t)\log\log\frac{1}{t(1-t)}\right)^{1/2} / q(t) < \infty$$

then, as $n \to \infty$,

$$\sup_{0<t<1}\sup_{\mathbf{s} \in I^d} |\hat{\alpha}_n(\mathbf{s},t)| / q(t) \xrightarrow{\mathcal{D}} \sup_{0,t<1}\sup_{\mathbf{s} \in I^d} |\Gamma_G(\mathbf{s},t)| / q(t), \tag{4}$$

(c) *if*
$$\int_0^1 \frac{(t(1-t))^{p/2}}{q(t)}dt < \infty, \quad 0 < p < \infty, \text{ then, as } n \to \infty$$
$$\int_0^1 \int_{I^d} \frac{|\hat{\alpha}_n(\mathbf{s},t)|^p}{q(t)} G(d\mathbf{s})dt \xrightarrow{\mathcal{D}} \int_0^1 \int_{I^d} \frac{|\Gamma_G(\mathbf{s},t)|^p}{q(t)} G(d\mathbf{s})dt. \tag{5}$$

5. Contiguous Alternatives. Consider now testing

H_0 : $X_i, 1 \le i \le n$, have the same distribution function $F(\mathbf{x}), \mathbf{x} = (x_1, \cdots, x_d) \in \mathbb{R}^d$,

versus

H_1 : X_i, $1 \le i \le n$, have the respective distribution functions $F_{in}(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^d$, all assumed to be absolutely continuous with respect to $F(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^d$, $d \ge 1$, where we assume (with $d = 1$, cf. Khmaladze and Parjanadze (1986))

$$\left(\frac{dF_{in}}{dF}(L^{-1}(\mathbf{u}))\right)^{1/2} = 1 + \frac{h_n(t, L^{-1}(\mathbf{u}))}{2n^{1/2}}, \frac{(i-1)}{n} < t \le \frac{1}{n}, i = 1, \cdots, n, \tag{6}$$

and that there exists a function $h \in L^2[0,1]^{d+1}$ such that

$$\int_{I^d} h(t, L^{-1}(\mathbf{u}))F(L^{-1}(d\mathbf{u})) = \int_{I^d} h(t, L^{-1}(\mathbf{u}))G(d\mathbf{u}) = 0 \tag{7}$$

for almost all $t \in [0,1]$, and, as $n \to \infty$, we have

$$\int_{I^{d+1}} \left(h_n(t, L^{-1}(\mathbf{u})) - h\left(\frac{[nt]}{n}, L^{-1}(\mathbf{u})\right)\right)^2 G(d\mathbf{u}) \to 0. \tag{8}$$

The sequence $F_{1n} \times \cdots \times F_{nn}$, $n = 1, 2, \cdots$, thus parametrized is contiguous to the sequence $F \times \cdots \times F$, via LeCam's first lemma.

The change-point alternative is accommodated by taking

$$h(t, L^{-1}(\mathbf{u})) = \mathbf{1}(t \ge \eta)h(L^{-1}(\mathbf{u})) \tag{9}$$

for some $h \in L^2[0,1]^d$ with respect to the measure $G(\mathbf{u}) = F(L^{-1}(\mathbf{u}))$, namely we consider

H_1^η: There is an $\eta \in (0,1)$ such that $X_1, \cdots, X_{[n\eta]}$ have the same distribution function $F(x)$, and $X_{[n\eta]+1}, \cdots, X_n$ have distributions $F_{([n\eta]+1)n}(\mathbf{x}), \cdots, F_{nn}(\mathbf{x})$ respectively.

Let

$$c(\mathbf{s},t) := \int_0^t \int_0^{\mathbf{s}} h(\tau, L^{-1}(\mathbf{u}))G(d\mathbf{u})d\tau = \int_0^t \int_{L^{-1}(0+)}^{L^{-1}(\mathbf{s})} h(\tau, \mathbf{x})F(d\mathbf{x})d\tau$$

and $d(\mathbf{s},t) = c(\mathbf{s},t) - tc(\mathbf{s},1)$. Then, by Theorem 1 and LeCam's third lemma, we have

THEOREM 2. *Assume H_1. Then, if q is as in (a) of Theorem 1, as $n \to \infty$,*

$$\hat{\alpha}_n(\mathbf{s},t)/q(t) \xrightarrow{\mathcal{D}} (\Gamma_G(\mathbf{s},t) + d(\mathbf{s},t))/q(t) \text{ in } D[0,1]^{d+1}.$$

Under H_1^η, $d(\mathbf{s},t)$ becomes

$$d_\eta(\mathbf{s},t) := (-t(1-\eta)\mathbf{1}(t<\eta) - (1-t)\eta\mathbf{1}(t \geq \eta)) \int_0^{\mathbf{s}} h(L^{-1}(\mathbf{u}))G(d\mathbf{u}).$$

3. Testing for Independence. We note that *if F is continuous and* $F(\mathbf{x}) = \prod_{j=1}^d F_{(j)}(x_j)$ for all $\mathbf{x} \in \mathbb{R}^d$, then $G(\mathbf{u}) = F(L^{-1}(\mathbf{u})) = \prod_{j=1}^d u_j = \lambda(\mathbf{u})$, the Lebesgue measure (uniform distribution) on I^d, $d \geq 1$. In this case $\Gamma_G = \Gamma_\lambda := \Gamma$ does not depend on $G = F(L^{-1})$, and covariance function in (2) reduces to

$$E\Gamma(\mathbf{s}_1,t_1)\Gamma(\mathbf{s}_2,t_2) = \left(\prod_{j=1}^d x_{1j} \wedge s_{2j} - \prod_{j=1}^d s_{1j}s_{2j}\right)(t_1 \wedge t_2 - t_1t_2).$$

Consider now testing

$H_0^{(1)}$: X_i, $1 \leq i \leq n$, have the same continuous distribution function $F(\mathbf{x}) = \prod_{j-=1}^d F_{(j)}(x_j)$, for all $\mathbf{x}$ in $\mathbb{R}^d$, $d \geq 2$,

versus

$H_1^{\eta_1}$: There is an $\eta_1 \in (0,1)$ such that $X_1, \cdots, X_{[n\eta_1]}$ have a continuous distribution function $F(\mathbf{x}) = \prod_{j=1}^d F_{(j)}(x_j)$ for all $\mathbf{x} \in \mathbb{R}^d$, and X_i, $i = [n\eta_1]+1, \cdots, n$, have a continuous distribution function $F(\mathbf{x}) \neq \prod_{j=1}^d F_{(j)}(x_j)$ for some $\mathbf{x} \in \mathbb{R}^d$, $d \geq 2$.

In this set-up then, tests based on (3), (4) and (5) will be consistent, distribution free nonparametric tests for $H_0^{(1)}$ *versus* the change-point alternative of $H_1^{\eta_1}$ that we change from random sampling with independent components to random sampling with nonindependent components.

For another approach to testing for independence in terms of empiricals, we refer to Hoeffding (1948), Blum, Kiefer and Rosenblatt (1961), Csörgő (1979), and Cotterill and Csörgő (1985). In our paper we will also study the empirical processes of these papers in weighted metrics and along the lines of Theorems 1 and 2. The question of determining critical values is, of course, crucial for applications of our results. A useful reference in this context is Romano (1989). In particular, his Example 5 is of special importance in this regard.

REFERENCES

BLUM, J. R., KIEFER, J. and ROSENBLATT, M. (1961). Distribution free tests of independence based on the sample distribution function. *Ann. Math. Statist.* **32**, 485–498.

COTTERILL, D. S. and CSÖRGŐ, M. (1985). On the limiting distribution of and critical values for the Hoeffding, Blum, Kiefer, Rosenblatt independence criterion. *Statist. Decisions* **3**, 1–48.

CSÖRGŐ, M. (1979). Strong approximations of the Hoeffding, Blum, Kiefer, Rosenblatt multivariate empirical process. *J. Multivariate Anal.* **9**, 84–100.

CSÖRGŐ, M. and HORVÁTH, L. (1988). A note on strong approximations of multivariate empirical processes. *Stochastic Process. Appl.* **28** 101–109.

DESHAYES, J. and PICARD, D. (1986). Off-line statistical analysis of changepoint models using non parametric and likelihood methods. In *Lecture Notes in Control and Information Sciences* (eds. M. Thoma and A. Wyner) **77**: Detection of Abrupt Changes in Signals and Dynamical Systems (eds. M. Basseville and A. Benveniste), 103–168. Springer-Verlag, Berlin.

HOEFFDING, W. (1948). A nonparametric test of independence *Ann. Math. Statist.* **19**, 546–557.

KHMALADZE, E. V. and PARJANADZE, A. M. (1986). Functional limit theorems for linear statistics of sequential ranks. *Probab. Theory Related Fields* **73**, 322–334.

ROMANO, J. P. (1989). Bootstrap and randomization tests of some nonparametric hypotheses. *Ann. Statist.* **17**, 141–159.

SZYSZKOWICZ, B. (1991). Weak convergence of empirical type processes under contiguous and changepoint alternatives. To appear in *Stochastic Process. Appl.*

DEPARTMENT OF MATHEMATICS & STATISTICS
CARLETON UNIVERSITY
OTTAWA, K1S 5B6, CANADA