4) *the 12 closed trihedral angles of the space* R^4 *generating by the hyperplanes* $\lambda_1 = \lambda_2$, $2\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 = 0$, $\lambda_1 + 2\lambda_2 + \lambda_3 + \lambda_4 = 0$ *and* $\alpha_3\lambda_3 + \alpha_4\lambda_4 = 0$, *where* $\alpha_3^2 + \alpha_4^2 = 1$,

5) *the 48 closed tetrahedral angles of the space* R^4 *generating by the hyperplanes* $\lambda_1 = \lambda_2$, $\lambda_2 = \lambda_3$, $\lambda_1 = \lambda_3$, $2\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 = 0$, $\lambda_1 + 2\lambda_2 + \lambda_3 + \lambda_4 = 0$, $\lambda_1 + \lambda_2 + 2\lambda_3 + \lambda_4 = 0$ *and* $\lambda_4 = 0$, *and, correspondingly,*

6) *the 24 closed tetrahedral angles of the space* R^4 *generating by the hyperplanes* $\lambda_1 = \lambda_2$, $\lambda_3 = \lambda_4$, $2\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 = 0$, $\lambda_1 + 2\lambda_2 + \lambda_3 + \lambda_4 = 0$ *and* $\lambda_3 + \lambda_4 = 0$.

When $j \neq 0$ *the complete system of base sets for the spectral densities* (23) *is formed by*

1) *the space* $\mathbf{R}^4$,

2) *the 2 closed half-spaces of the space* R^4 *generated the hyperplane* $2\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 = j\omega$,

3) *the 4 closed dihedral angles of the space* R^4 *generated the hyperplanes* $2\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 = j\omega$ *and* $\lambda_3 = \lambda_4$,

4) *the 6 closed dihedral angles of the space* R^4 *generated the hyperplanes* $\lambda_1 = \lambda_2$, $2\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 = j\omega$ *and* $\lambda_1 + 2\lambda_2 + \lambda_3 + \lambda_4 = j\omega$,

5) *the 24 closed trihedral angles of the space* R^4 *generated the hyperplanes* $\lambda_1 = \lambda_2$, $\lambda_2 = \lambda_3$, $\lambda_1 = \lambda_3$, $2\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 = j\omega$, $\lambda_1 + 2\lambda_2 + \lambda_3 + \lambda_4 = j\omega$, $\lambda_1 + \lambda_2 + 2\lambda_3 + \lambda_4 = j\omega$, *and, correspondingly,*

6) *the 12 closed trihedral angles of the space* R^4 *generated the hyperplanes* $\lambda_1 = \lambda_2$, $\lambda_3 = \lambda_4$, $2\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 = j\omega$, $\lambda_1 + 2\lambda_2 + \lambda_3 + \lambda_4 = j\omega$.

6. Till now we assumed that our stochastic process $\xi(t)$ is periodically nonstationary. As for a stationary stochastic process $\xi(t)$, $t \in \mathbf{R}$, the symmetry properties of its cross spectral densities are the same as those described above for spectral densities corresponding to the case $j = 0$. Thus, the case of a stationary process $\xi(t)$ needs no special consideration.

The symmetry properties found for all spectral densities considered above allow one to shorten the calculations necessary for their statistical estimation.

REFERENCES

[1] V. G. ALEKSEEV, *On the symmetry properties of higher-order spectral densities for stationary stochastic processes*, Math. Notes, 41 (1987), pp. 758–763. (In Russian.)

[2] ———, *On the symmetry properties of higher-order spectral densities for stationary and periodically nonstationary stochastic processes*, Problems of Inform. Transmission, 23 (1987), pp. 48–53. (In Russian.)

[3] N. K. BARY, *A Treatise on Trigonometric Series*, Pergamon Press, New York, 1964.

NONPARAMETRIC CHANGE-POINT ESTIMATION FOR DATA FROM AN ERGODIC SEQUENCE*

E. CARLSTEIN† AND S. LELE‡

Abstract. In the framework of the series scheme we assume that an observations sequence $\{X_i^n,\ 1 \leqq i \leqq n\}$ is such that $X_i^n = U_i I\,(1 \leqq i \leqq [\theta n]) + V_i I([\theta n] + 1 \leqq i \leqq n)$, where (U_i, V_i) is a stationary ergodic sequence the marginal distributions of which are different, and θ is a change-point in the probabilistic characteristics such that $\theta \in (0;1)$. The main result of this paper is the proof

*Received by the editors November 8, 1993. This research was supported by National Scince Foundation grant #DMS-8902973

†University of North Carolina at Chapel Hill, CB#3260, Phillips Hall, Chapel Hill, N.C. 27599–3260.

‡Johns Hopkins University, Visiting the University of North Carolina at Chapel Hill, August 1989.

of the fact that the sequence $(\theta n)_{n \geqq 1}$ of nonparametric estimations constructed here is consistent $(\theta_n \to \theta)$.

Key words. nonparametric estimation of a change-point in the probabilistic characteristics, consistency of estimations

1. Introduction. Suppose we observe a sequence of random variables $X_1^n, X_2^n, \ldots, X_n^n$, where

X_i^n has marginal distribution F for $1 \leqq i \leqq [\theta n]$,

X_i^n has marginal distribution G for $[\theta n] + 1 \leqq i \leqq n$;

here $[y]$ denotes the greatest integer not exceeding y. The unknown parameter $\theta \in (0, 1)$ is the *change-point* to be estimated. The purpose of this note is simply to show that θ can be consistently estimated in a fully nonparametric scenario:

(i) No knowledge of F or G is required. There are *no parametric assumptions* (e.g., normality) and *no regularity conditions* (e.g., continuity) on F or G;

(ii) No restrictions are imposed on the *strength of serial dependence* in $\{X_i^n\colon\ 1 \leqq i \leqq n\}$, beyond the minimal condition of ergodicity; there are *no assumptions about the dependence mechanism* (e.g., autoregression), and there are no "*mixing*" conditions;

(iii) No prior restrictions on θ are needed.

In fact, a whole *class* of such fully nonparametric estimators will be presented.

A huge amount of work has been done on change-point estimation in general (see [13] for an annotated bibliography and [11] for a recent extensive review); the importance of the *nonparametric* approach in particular is well established (see [5] for a review). Most of the *nonparametric* methods assume *independence* in $\{X_i^n\colon\ 1 \leqq i \leqq n\}$; and most of them assume *prior knowledge* about how F and G differ (e.g., in their means, medians, or other measure of level). However, there has recently been a trend towards eliminating prior assumptions on F and G, and towards allowing for serial dependence in $\{X_i^n\colon\ 1 \leqq i \leqq n\}$ (see [1]–[10] and [12]).

Within the nonparametric context, it is certainly desirable to minimize the assumptions on F and G. Since change-point data is inherently time-sequential, it is natural and practical to allow for serial dependence. It is desirable to minimize any restriction on the *strength* of this serial dependence, because the *joint* distribution structure of $\{X_i^n\colon\ 1 \leqq i \leqq n\}$ is much more obscure than the *marginal* distribution structure (i.e., F and G), so it is unrealistic to assume knowledge about the former when the latter is completely unknown. Moreover, the usual "*mixing*" assumptions are impossible to check for a given set of data. Our main consistency result (in § 3) shows how far we can relax the restrictions on F, G, and the dependence structure.

Let us briefly contrast our fully nonparametric approach (i.e., (i), (ii), (iii) above) with the related works ([1]–[10] and [12]) which have recently appeared in the literature. In [4] and [5], it is assumed that the X_i^ns are *independent* and that F and G are *continuous*. In [2], [6], [8], and [9], it is assumed that F and G are both *discrete* with *finite support*, and that the sequence of X_i^ns satisfies a *strong-mixing* condition; it is also assumed that θ is in the *known* interval $[\alpha, \beta]$, where $0 < \alpha < \beta < 1$. In [7], either one of two possible scenarios is required: F and G are both *discrete* with *finite support*, and the X_i^ns are *strong-mixing*; or, F and G are both *continuous* with *support in* $[0, 1]$ and satisfy a *Lipschitz condition*, and the X_i^ns are *ψ-mixing*. In [1], it is assumed that the X_i^ns are *independent* (or possibly *m-dependent*) and that F and G are *continuous*; it is also assumed that θ is in the *known* interval $[\alpha, \beta]$, $0 < \alpha < \beta < 1$. In [3] and [10], the X_i^ns are again assumed to be *independent*. In [12], it is assumed that the X_i^ns arise from a *Gaussian process*, with F and G sharing the *same mean*. We shall impose *none* of these assumptions.

One final point of comparison between our approach and the works cited above: in each of the references [2], [4]–[10], and [12], a certain "norm" is used to calculate the basic statistic (this notion of a "norm" will be made precise in § 2). Specifically, a *Kolmogorov–Smirnov norm* is used in [4], [5], [7], [10], [12]; and a *Cramér–von Mises norm* is used in [2], [6], [8],

[9], [12]. Our method is based on a general "*Mean-Dominant norm*" (defined in § 2), which includes as special cases the Kolmogorov–Smirnov norm, the Cramér–von Mises norm, as well as many other norms. Thus, our approach actually provides a whole *class* of consistent nonparametric estimators; moreover, our general formulation allows us to simultaneously analyze estimators based on *all* these norms with one unified theoretical argument (see §§ 2 and 4).

2. The estimator. The estimator is constructed from the *pre-t empirical c.d.f. (cumulative distribution function)*:

$$ {}_t h^n(x) := \sum_{j=1}^{nt} \frac{I\{X_j^n \leqq x\}}{nt}, $$

and the *post-t empirical c.d.f.*:

$$ h_t^n(x) := \sum_{j=nt+1}^{n} \frac{I\{X_j^n \leqq x\}}{n(1-t)}. $$

The index t corresponds to *possible* values of the change-point estimator; we will only need to consider $t \in T_n$, where

$$ T_n := \left\{ \frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n} \right\} \bigcap [\alpha_n, 1-\alpha_n], $$

and $\{\alpha_n\colon\ n \geqq 1\}$ is any deterministic sequence satisfying $\alpha_n \downarrow 0$ and $n\alpha_n \uparrow \infty$ as $n \to \infty$.

For fixed $t \in T_n$, compute the differences between the pre-t and post-t empirical c.d.f.'s *at the sample observations*, i.e.,

$$ d_{ni}^t := \left| {}_t h^n(X_i^n) - h_t^n(X_i^n) \right|, \qquad 1 \leqq i \leqq n. $$

These n differences are now combined via a "Mean-Dominant norm" $S_n\colon\ \mathbf{R}^n \mapsto \mathbf{R}$, yielding the criterion function

$$ D_n(t) := t\,(1-t)\, S_n(d_{n1}^t, d_{n2}^t, \dots, d_{nn}^t). $$

Our fully nonparametric estimator θ_n is then defined to be:

$$ \theta_n \in T_n \quad \text{such that} \quad D_n(\theta_n) = \max_{t \in T_n} D_n(t). $$

The "Mean-Dominant norm" $S_n(\cdot,\cdot,\dots,\cdot)$ is any function satisfying the following natural conditions (whenever the arguments d_i and d_i' are all nonnegative):

(1) [Symmetry] $S_n(\cdot,\cdot,\dots,\cdot)$ is symmetric in n arguments;
(2) [Homogeneity] $S_n(cd_1, cd_2, \dots, cd_n) = cS_n(d_1, d_2, \dots, d_n)$ whenever $c \geqq 0$;
(3) [Triangle Inequality]

$$ S_n(d_1+d_1', d_2+d_2', \dots, d_n+d_n') \leqq S_n(d_1, d_2, \dots, d_n) + S_n(d_1', d_2', \dots, d_n'); $$

(4) [Identity] $S_n(1, 1, \dots, 1) = 1$;
(5) [Monotonicity] $S_n(d_1, d_2, \dots, d_n) \leqq S_n(d_1', d_2', \dots, d_n')$ whenever $d_i \leqq d_i'$ $\forall i$;
(6) [Mean-Dominance] $S_n(d_1, d_2, \dots, d_n) \geqq \sum_{1 \leqq i \leqq n} d_i/n$.

Special cases of the "Mean-Dominant norm" include the *Kolmogorov–Smirnov* norm

$$ S_n^{\mathrm{KS}}(d_1, d_2, \dots, d_n) := \sup_{1 \leqq i \leqq n} \{d_i\}, $$

the *Cramér–von Mises* norm

$$ S_n^{\mathrm{CM}}(d_1, d_2, \dots, d_n) := \left(\sum_{1 \leqq i \leqq n} \frac{d_i^2}{n} \right)^{1/2}, $$

and the *arithmetic mean* norm

$$S_n^{\mathrm{am}}(d_1, d_2, \ldots, d_n) := \sum_{1 \leqq i \leqq n} \frac{d_i}{n}.$$

The intuition behind our method is as follows. The c.d.f.'s ${}_th^n(x)$ and $h_t^n(x)$ are empirical approximations of the unknown distributions

$${}_th(x) := I\{t \leqq \theta\}\, F(x) + I\{t > \theta\} \Big(\theta F(x) + (t-\theta)\, G(x)\Big)/t$$

and

$$h_t(x) := I\{t \leqq \theta\} \Big((\theta - t)\, F(x) + (1-\theta)\, G(x)\Big)/(1-t) + I\{t > \theta\}\, G(x),$$

respectively. Therefore, the differences d_{ni}^t are empirical approximations of

$$\delta_{ni}^t := \Big|{}_th(X_i^n) - h_t(X_i^n)\Big| = \delta_{ni}^\theta \Big(I\{t \leqq \theta\}\,(1-\theta)/(1-t) + I\{t > \theta\}\theta/t\Big), \qquad 1 \leqq i \leqq n.$$

And, the criterion function $D_n(t)$ is an empirical approximation of the corresponding function

$$\Delta_n(t) := t\,(1-t)\, S_n(\delta_{n1}^t, \delta_{n2}^t, \ldots, \delta_{nn}^t).$$

Now, by S_n's homogeneity, we have

$$\Delta_n(t) := \rho(t)\, S_n(\delta_{n1}^\theta, \delta_{n2}^\theta, \ldots, \delta_{nn}^\theta),$$

where

$$\rho(t) := I\{t \leqq \theta\}\, t\,(1-\theta) + I\{t > \theta\}\,(1-t)\,\theta.$$

Notice that the maximizer of $\Delta_n(t)$ over $t \in (0,1)$ is *precisely* at $t = \theta$. Thus, the maximizer of the analogous sample-based criterion function $D_n(t)$ is a reasonable estimator of θ. This logic applies for *any* "Mean-Dominant norm" S_n.

3. Main result. The data $\{X_i^n\colon\ 1 \leqq i \leqq n\}$ are embedded in a stationary ergodic sequence

$$\Big\{(U_i, V_i)\colon\ -\infty < i < +\infty\Big\},$$

where U_i has marginal distribution F, and V_i has marginal distribution G. Specifically, the data arise as

$$X_i^n = U_i I\Big\{1 \leqq i \leqq [\theta n]\Big\} + V_i I\Big\{[\theta n] + 1 \leqq i \leqq n\Big\}.$$

There will be no further constraints on the serial dependence structure of $\{X_i^n\colon\ 1 \leqq i \leqq n\}$. The only assumption on the unknown marginal distributions is simply that $F \neq G$. The unknown change-point parameter $\theta \in (0,1)$ is unrestricted. In this scenario, our fully nonparametric estimator θ_n (defined in § 2) is consistent.

THEOREM. *$\theta_n \xrightarrow{\mathbf{P}} \theta$ as $n \to \infty$.*

See § 4 for a proof of this result.

4. Proof. We begin by presenting three preliminary Lemmas which will be used in proving the Theorem.

LEMMA 1. *Denote*

$$\mu_F := \int_{-\infty}^{\infty} \Big|F(x) - G(x)\Big|\, dF(x),$$

$$\mu_G := \int_{-\infty}^{\infty} \Big|F(x) - G(x)\Big|\, dG(x),$$

$$\mu := \theta\mu_F + (1-\theta)\,\mu_G.$$

Then $\mu > 0$.

Proof. By assumption we have

$$\Lambda := \left\{ x \in \mathbf{R}\colon \ \left| F(x) - G(x) \right| > 0 \right\} \neq \varnothing.$$

It suffices to show that either $\int_\Lambda dF(x) > 0$ or $\int_\Lambda dG(x) > 0$. The case where Λ contains a discontinuity point of F or G is trivial, so we will now presume that F and G are continuous at each $x \in \Lambda$.

Choose $x_0 \in \Lambda$ with (say) $F(x_0) > G(x_0)$. Then

$$\sigma := \left\{ y \in (-\infty, x_0)\colon \ F(x) > G(x) \ \forall\, x \in (y, x_0] \right\}$$

is nonempty, by continuity. Denote $y_0 := \inf\{y \in \sigma\}$. If $y_0 = -\infty$, then $(-\infty, x_0] \subseteq \Lambda$ and, therefore,

$$\int_\Lambda dF(x) \geqq F(x_0) > G(x_0) \geqq 0.$$

If $y_0 > -\infty$, then $F(y_0) \leqq G(y_0)$ and also $(y_0, x_0] \subseteq \Lambda$, yielding

$$\int_\Lambda dF(x) \geqq F(x_0) - F(y_0) \geqq F(x_0) - G(y_0) > G(x_0) - G(y_0) \geqq 0.$$

Lemma 1 is proved.

LEMMA 2. *Let $\{Y_i\colon \ -\infty < i < +\infty\}$ be a stationary ergodic sequence, where Q is the marginal distribution of Y_i. Let $\{a_n\colon \ n \geqq 1\}$ be a deterministic sequence satisfying $a_n \uparrow \infty$ as $n \to \infty$, and let $\{N_n\colon \ n \geqq 1\}$ be integer-valued random variables for which $N_n \geqq a_n$ for any n. Define for $m_2 \geqq m_1 + 1$,*

$$W_Q(m_1, m_2) := \sup_{-\infty < y < \infty} \left| \sum_{i=m_1+1}^{m_2} \frac{I\{Y_i \leqq y\}}{m_2 - m_1} - Q(y) \right|.$$

Then $W_Q(0, N_n) \xrightarrow{\text{a.s.}} 0$ as $n \to \infty$.

Proof. Follows directly from the theorem in [14].

LEMMA 3. *Denote $A_n := T_n \cap \{[\alpha_n, \theta - \alpha_n] \cup (\theta + \alpha_n, 1 - \alpha_n]\}$. Let $\{R_n\colon \ n \geqq 1\}$ be a (possibly random) sequence with $R_n \equiv r_n((U_1, V_1), (U_2, V_2), \dots, (U_n, V_n))$ for a deterministic function $r_n\colon \ \mathbf{R}^{2n} \mapsto A_n$. Then*

$$\left| \Delta_n(R_n) - D_n(R_n) \right| \xrightarrow{\mathbf{P}} 0 \qquad as \quad n \to \infty.$$

Proof. Denote $e_{ni} := {}_{ni}H + H_{ni}$, where

$$H_{ni} := \left| h^n_{R_n}(X^n_i) - h_{R_n}(X^n_i) \right|, \qquad {}_{ni}H := \left| {}_{R_n}h^n(X^n_i) - {}_{R_n}h(X^n_i) \right|,$$

so that $d^{R_n}_{ni} \leqq e_{ni} + \delta^{R_n}_{ni}$ and, therefore:

$$\begin{aligned} \Gamma_n := D_n(R_n) - \Delta_n(R_n) &\leqq S_n(e_{n1}, e_{n2}, \dots, e_{nn}) \\ &\leqq S_n({}_{n1}H, {}_{n2}H, \dots, {}_{nn}H) + S_n(H_{n1}, H_{n2}, \dots, H_{nn}) \end{aligned}$$

by (5) and (3). The same bound holds for $-\Gamma_n$ and hence for $|\Gamma_n|$. Now,

$$\begin{aligned} H_{ni} = \Bigg| & I\{R_n > \theta\} \left(\sum_{j=nR_n+1}^{n} \frac{I\{X^n_j \leqq X^n_i\}}{n - nR_n} - G(X^n_i) \right) \\ & + I\{R_n \leqq \theta\} \left(\left(\sum_{j=nR_n+1}^{[\theta n]} \frac{I\{X^n_j \leqq X^n_i\}}{[\theta n] - nR_n} - F(X^n_i) \right) \frac{[\theta n] - nR_n}{n - nR_n} \right. \end{aligned}$$

$$+ F(X_i^n)\left(\frac{[\theta n] - nR_n}{n - nR_n} - \frac{\theta - R_n}{1 - R_n}\right)$$

$$+ \left(\sum_{j=[\theta n]+1}^{n} \frac{I\{X_j^n \leqq X_i^n\}}{n - [\theta n]} - G(X_i^n)\right) \frac{n - [\theta n]}{n - nR_n}$$

$$+ G(X_i^n)\left(\frac{n - [\theta n]}{n - nR_n} - \frac{1-\theta}{1 - R_n}\right)\Bigg)\Bigg| \leqq I\{R_n > \theta\}\, W_G(nR_n, n)$$

$$+ I\Big\{R_n \leqq \theta\Big\}\Big(W_F\Big(nR_n, [\theta n]\Big) + W_G\Big([\theta_n], n\Big) + (2/n)(1 - R_n)\Big)$$

$$\leqq W_G\Big(I\{R_n > \theta\}\, n\, R_n + I\{R_n \leqq \theta\}\, [\theta n], n\Big)$$

$$+ W_F\Big(I\{R_n \leqq \theta\}\, n\, R_n,\, [\theta n]\Big) + (2/n)\,(1-\theta) =: H_n.$$

(In $W_F(\cdot)$, V_i plays the role of Y_i; in $W_G(\cdot)$, V_i plays the role of Y_i.) An analogous bound (say, ${}_nH$) can be obtained for ${}_{ni}H$). Since H_n and ${}_nH$ do not depend on i, we have $|\Gamma_n| \leqq {}_nH + H_n$, by (5), (2), and (4).

We will now show that $H_n \xrightarrow{\mathbf{P}} 0$; a similar argument applies to ${}_nH$. We deal explicitly with the second summand in H_n; the first summand can be handled analogously to the second, and the third summand goes to zero deterministically. Denote $Z_i := (U_i, V_i)$ and $\mathcal{Z}_n^k := (Z_{k+1}, Z_{k+2}, \ldots, Z_{k+n})$. Note that:

$$W_F\Big(I\Big\{r_n(\mathcal{Z}_n^0) \leqq \theta\Big\}\, nr_n(\mathcal{Z}_n^0), [\theta n]\Big)$$

$$\stackrel{\mathcal{D}}{=} W_F\Big(I\Big\{r_n(\mathcal{Z}_n^{-[\theta n]-1}) \leqq \theta\Big\}\, nr_n(\mathcal{Z}_n^{-[\theta n]-1}) - [\theta n] - 1, -1\Big)$$

$$= W_F\Big(0, [\theta n] - I\Big\{r_n\,(\mathcal{Z}_n^{-[\theta n]-1}) \leqq \theta\Big\}\, nr_n(\mathcal{Z}_n^{-[\theta n]-1})\Big),$$

where the last expression uses $Y_i \equiv U_{-i}$. We can now apply Lemma 2, since $r_n(\cdot) \in A_n$ and so:

$$[\theta n] - I\Big\{r_n(\cdot) \leqq \theta\Big\}\, nr_n(\cdot) \geqq [\theta n] - n\,(\theta - \alpha_n) \geqq (n\alpha_n - 1) \uparrow \infty.$$

This establishes $W_F(I\{R_n \leqq \theta\}nR_n, [\theta n]) \xrightarrow{\mathbf{P}} 0$.

The following notation and definitions will be needed to prove the theorem. The random variable $\tilde{\theta}_n$ is defined as follows:

if $\theta_n \in A_n$, then let $\tilde{\theta}_n = \theta_n$;

if $\theta_n \notin A_n$, then let $\tilde{\theta}_n \in A_n$ satisfy $D_n(\tilde{\theta}_n) = \max_{t \in A_n} D_n(t)$.

Note that, in either case, we have $\tilde{\theta}_n \in A_n$ and $D_n(\tilde{\theta}_n) = \max_{t \in A_n} D_n(t)$. Also define the nonrandom entity

$$t_n := I\Big\{\rho\Big([(\theta - \alpha_n)\, n]/n\Big) \geqq \rho\Big(([(\theta + \alpha_n)\, n] + 1)/n\Big)\Big\}\Big[(\theta - \alpha_n)\, n\Big]/n$$

$$+ I\Big\{\rho\Big([(\theta - \alpha_n)\, n]/n\Big) < \rho\Big(([(\theta + \alpha_n)\, n] + 1)/n\Big)\Big\}\Big([(\theta + \alpha_n)\, n] + 1\Big)\Big/n,$$

which satisfies $t_n \in A_n$ and $\Delta_n(t_n) = \max_{t \in A_n} \Delta_n(t)$.

In order to prove the theorem, it suffices to establish that $\tilde{\theta}_n \xrightarrow{\mathbf{P}} \theta$ as $n \to \infty$. To see that this is sufficient, write

$$\mathbf{P}\Big\{|\theta_n - \theta| > \varepsilon\Big\} = \mathbf{P}\Big\{|\theta_n - \theta| > \varepsilon,\ \theta_n \in A_n\Big\} + \mathbf{P}\Big\{|\theta_n - \theta| > \varepsilon,\ \theta_n \notin A_n\Big\}.$$

The latter probability is zero for n sufficiently large; the other probability is bounded by $\mathbf{P}\{|\tilde{\theta}_n - \theta| > \varepsilon\}$.

To establish $\tilde{\theta} \xrightarrow{\mathbf{P}} \theta$, denote $\delta_n := \sum_{1 \leqq i \leqq n} \delta_{ni}^{\theta}/n$ and $\bar{\theta} := \min\{\theta, 1-\theta\}$, and notice that:

$$\Psi_n := \Delta_n(\theta) - \Delta_n(\tilde{\theta}_n) = \Big(\rho(\theta) - \rho(\tilde{\theta}_n)\Big)\, S_n(\delta_{n1}^{\theta}, \delta_{n2}^{\theta}, \ldots, \delta_{nn}^{\theta})$$
$$\geqq \Big(I\{\tilde{\theta}_n \leqq \theta\}\,(\theta - \tilde{\theta}_n)\,(1-\theta) + I\{\tilde{\theta}_n > \theta\}\,(\tilde{\theta}_n - \theta)\,\theta\Big)\,\delta_n \geqq |\tilde{\theta}_n - \theta|\,\bar{\theta}\delta_n,$$

where the first inequality follows from (6). Therefore,

$$\mathbf{P}\left\{|\tilde{\theta}_n - \theta| > \varepsilon\right\} \leqq \mathbf{P}\{\Psi_n > \varepsilon\bar{\theta}\,\delta_n,\ \delta_n \geqq \nu\} + \mathbf{P}\{\Psi_n > \varepsilon\bar{\theta}\,\delta_n,\ \delta_n < \nu\}$$
$$\leqq \mathbf{P}\{\Psi_n > \varepsilon\bar{\theta}\,\nu\} + \mathbf{P}\{\delta_n < \nu\},$$

where $\nu := \mu/2 > 0$ by Lemma 1. To deal with the latter probability, write

$$\delta_n = \frac{[\theta n]\,\underline{\delta}_n + (n - [\theta n])\,\overline{\delta}_n}{n},$$

where

$$\underline{\delta}_n := \sum_{1 \leqq i \leqq [\theta n]} \frac{\delta_{ni}^{\theta}}{[\theta n]}, \qquad \overline{\delta}_n := \sum_{[\theta n]+1 \leqq i \leqq n} \frac{\delta_{ni}^{\theta}}{n - [\theta n]}.$$

Note that

$$\{\delta_n < \nu\} \Longrightarrow \Big\{\nu < |\delta_n - \mu| \leqq |\underline{\delta}_n - \mu_F| + |\overline{\delta}_n - \mu_G| + 2\Big|\theta - [\theta n]/n\Big|\Big\}.$$

Since the term in the last modulus goes to zero deterministically, we only need to consider $|\overline{\delta}_n - \mu_G|$ (the remaining term is handled analogously). Now

$$\overline{\delta}_n^{*} := \sum_{1 \leqq i \leqq n - [\theta n]} \frac{|F(V_i) - G(V_i)|}{n - [\theta n]} \xrightarrow{\mathbf{P}} \mu_G$$

by the ergodic theorem, and $\overline{\delta}_n \overset{\mathcal{D}}{=} \overline{\delta}_n^{*}$, so that $|\overline{\delta}_n - \mu_G| \xrightarrow{\mathbf{P}} 0$.

To prove the theorem, it now suffices to show that $\Psi_n \xrightarrow{\mathbf{P}} 0$. We have

$$\Psi_n \leqq \Big|\Delta_n(\theta) - \Delta_n(t_n)\Big| + \Big|\Delta_n(t_n) - D_n(\tilde{\theta}_n)\Big| + \Big|D_n(\tilde{\theta}_n) - \Delta_n(\tilde{\theta}_n)\Big|,$$

where the first modulus is deterministically dominated (using (5) and (4)) by:

$$\rho(\theta) - \rho(t_n) = I\{t_n \leqq \theta\}\,(\theta - t_n)\,(1-\theta) + I\{t_n > \theta\}\,(t_n - \theta)\,\theta \leqq \alpha_n + n^{-1},$$

and where the second modulus is dominated by

$$\Big|D_n(t_n) - \Delta_n(t_n)\Big| + \Big|D_n(\tilde{\theta}_n) - \Delta_n(\tilde{\theta}_n)\Big|$$

because either

$$\Delta_n(t_n) \geqq D_n(\tilde{\theta}_n) \geqq D_n(t_n)$$

or

$$D_n(\tilde{\theta}_n) \geqq \Delta_n(t_n) \geqq \Delta_n(\tilde{\theta}_n)$$

holds. Thus, applying Lemma 3 with $R_n \equiv t_n$ and with $R_n \equiv \tilde{\theta}_n$ completes proof of the Theorem.

REFERENCES

[1] D. Asatryan and I. Safaryan, *Nonparametric methods for detecting changes in the properties of random sequences*, In: Detection of Changes in Random Processes, Ed. by L. Telksnys. New York, 1986, pp. 1–13.

[2] B. BRODSKII AND B. DARKHOVSKII, *The a posteriori method of detecting the disruption of a random field*, In: Detection of Changes in Random Processes, Ed. by L. Telksnys, New York, 1986, pp. 32–38.

[3] E. CARLSTEIN, *Nonparametric change-point estimation*, Ann. Statist., 16 (1988), pp. 188–197.

[4] M. CSÖRGÖ AND L. HORVÁTH, *Nonparametric tests for the changepoint problem*, J. Statist. Planning Inference, 17 (1987), pp. 1–9.

[5] ———, *Nonparametric methods for changepoint problems*, In: Handbook of Statistics, vol. 7, Ed. by P. Krishnaiah and C. R. Rao, Elsevier, Netherlands, 1988, pp. 403–425.

[6] B. DARKHOVSKII, *On two estimation problems for times of change of the probabilistic characteristics of a random sequence*, Theory Probab. Appl., 29 (1984), pp. 478–487.

[7] ———, *A nonparametric method of estimating intervals of homogeneity for a random sequence*, Theory Probab. Appl., 30 (1985), pp. 845–849.

[8] ———, *A general method for estimating the instant of change in the probabilistic characteristics of a random sequence*, In: Detection of Changes in Random Processes, Ed. by L. Telksnys. New York, 1986, pp. 47–52.

[9] B. DARKHOVSKII AND B. BRODSKII, *A posteriori detection of the "disorder" time of a random sequence*, Theory Probab. Appl., 25 (1980), pp. 624–628.

[10] J. DESHAYES AND D. PICARD, *Convergence de processus à double indice: application aux tests de rupture dans un modèle*, C. R. Acad. Sci. Paris, vol. 292 (1981), pp. 449–452.

[11] P. KRISHNAIAH AND B. MIAO, *Review about estimation of change points*, In: Handbook of Statistics, vol. 7, Ed. by P. Krishnaiah and C. R. Rao, Elsevier, Netherlands, 1988, pp. 375–402.

[12] D. PICARD, *Testing and estimating change-points in time series*, Adv. Appl. Probab., 17 (1985), pp. 841–867.

[13] S. SHABAN, *Change-point problem and two-phase regression: an annotated bibliography*, International Statist. Review, 48 (1980), pp. 83–93.

[14] H. TUCKER, *A generalization of the Glivenko–Cantelli theorem*, Ann. Math Statist., 30 (1959), pp. 828–830.

LAWS OF LARGE NUMBERS IN BANACH SPACES OF TYPE (F, F_1)*

A. N. CHUPRUNOV†

(Translated by M. V. Khatuntseva)

Abstract. This paper gives strong and weak laws of large numbers for random elements with values in Banach spaces of type (F, F_1). The known laws of large numbers in Banach spaces of stable and Rademacher type p are special cases of these results. Characterizations of spaces of type (F, F_1) are given in terms of these laws.

Key words. space of type (F, F_1), law of large numbers, random element

In this paper we obtain strong and weak laws of large numbers for random elements with values in Banach spaces of type (F, F_1). Normalizations of a more general type are also considered.

Let $\{a_n,\ n \in \mathbf{N}\}$ be a numerical sequence such that $a_n > 0$ $(n \in \mathbf{N})$ and $a_n \to \infty$ $(n \to \infty)$. Let X_n be random elements (r.e.'s) with values in Banach space (B-space) B. The strong law of large numbers (SLLN) (the weak law of large numbers (WLLN) respectively) is valid for $(\{X_n\}, \{a_n\})$ if

$$\frac{1}{a_n} \sum_{k=1}^{n} X_k \longrightarrow 0, \qquad n \to \infty,$$

*Received by the editors December 26, 1989.

†Kazan' State University, Kazan', Tatarstan, Russia.