Теорія Ймовір. та Матем. Статист.
Вип. 70, 2004

Theor. Probability and Math. Statist.
No. 70, 2005, Pages 83–92
S 0094-9000(05)00642-3
Article electronically published on August 5, 2005

IMPROVED ESTIMATORS FOR MOMENTS CONSTRUCTED FROM OBSERVATIONS OF A MIXTURE

UDC 519.21

R. MAĬBORODA AND O. KUBAĬCHUK

Abstract. Procedures for improving weighted empirical distribution functions constructed from mixtures with varying concentrations are considered. The procedures are such that the estimators of moments of the mixture components constructed from weighted empirical distribution functions have specified properties (say, estimators of the variance must not be negative). We prove that the moment estimators constructed from improved weighted empirical distribution functions have the same asymptotic behavior as those constructed from the original weighted empirical distribution functions.

1. Introduction

The model of a mixture with varying concentrations is often used to describe statistical data. A sample $\Xi_N = \{\xi_{j:N}, j = 1, \dots, N\}$ in this model (see [1]) consists of independent random variables $\xi_{j:N}$ with distributions

$$\mathsf{P}\{\xi_{j:N} < x\} = \sum_{m=1}^{M} w_{j:N}^{m} H_m(x) \tag{1}$$

where H_m is the distribution function of the component m in the mixture, and $w_{j:N}$ is the concentration of the component m in the mixture for the observation j.

We consider estimators for moments of distributions of components, namely

$$\bar{g}_k := \int g(x)\, H_k(dx) \tag{2}$$

where the function g is fixed, the concentrations of components are known, and the distributions H_k are unknown.

It is proposed in [2] to use the integral of g with respect to the weighted empirical distribution function, denoted by $\hat{F}_N$, as an estimator of $\bar{g}_k$:

$$\hat{F}_N(x, a) := \frac{1}{N} \sum_{j=1}^{N} a_{j:N}\, \mathbf{1}\{\xi_{j:N} < x\} \tag{3}$$

where $a_{j:N}$ are nonrandom weighted coefficients chosen so that $\hat{F}_N(x, a)$ be a nice estimator of $H_m(x)$. (For example, if $a = a^m$ are the coefficients introduced in [1], then $\hat{F}_N(x, a)$ is the minimax estimator in the class of unbiased estimators for the quadratic

2000 *Mathematics Subject Classification.* Primary 62G05; Secondary 62G20.

risk function.) Thus the estimator for $\bar{g}$ becomes of the form

$$\hat{g}_{k:N} := \int g(x)\,\hat{F}_N(dx,a) = \frac{1}{N}\sum_{j=1}^{N} a_{j:N} g(\xi_{j:N}). \tag{4}$$

Unfortunately the weighted coefficients $a_{j:N}$ should necessarily be negative for some j if one requires nice properties (unbiasedness, say) of the estimator $\hat{F}_N$. As a rule $\hat{F}_N$ is not a probability distribution function in this case. This leads to some problems, say the estimator for the second moment ($g(x) = x^2$) can be negative, etc.

To avoid the problems mentioned above, a method to improve the weighted empirical distribution function is proposed in [3]. For example, one can put

$$\tilde{F}_N^+(x,a) := \min\left(\sup_{y<x} \hat{F}_N(y,a), 1\right). \tag{5}$$

It is easy to see that $\tilde{F}_N^+(x,a)$ is a distribution function of a probability measure. Similarly, one can consider

$$\tilde{g}_{k:N}^+ := \int g(x)\,\tilde{F}_N^+(dx,a) \tag{6}$$

as an estimator for $\bar{g}$.

The aim of this paper is to study the asymptotic behavior of estimators (6) and similar estimators as the size of the sample is increasing. The main result is that, under certain assumptions, the estimators are normal with the same dispersion coefficient as that in the case of estimators (4). This, in particular, means that the asymptotic behavior of improved estimators is not worse than that of the estimators (4).

2. Improved weighted empirical distribution functions

Besides the estimator $\tilde{F}_N^+$ defined by (5) we consider some other improved estimators of the weighted empirical distribution functions. We assume that all the estimators defined below are continuous from the left and this corresponds to our definition of the distribution function F_ξ of a random variable ξ:

$$F_\xi(x) = \mathsf{P}\{\xi < x\}.$$

However, there are some cases where it is convenient to deal with an improved function continuous from the right. In such a case we "improve" an estimator by considering the operator

$$L[f](x) = \lim_{y\uparrow x} f(y)$$

that substitutes the left limit of the function f for its value at any point of discontinuity of f. Note however that this "improvement" has no sense from the point of view of the discussion that follows, since all the weighted empirical distribution functions considered below are of the form

$$\frac{1}{N}\sum_{j=1}^{N} b_j\,\mathbf{1}\{\xi_{j:N} < x\}.$$

To use such estimators one only needs to evaluate the coefficients b_j (algorithms for the evaluation of the coefficients are described in [3]).

Let

$$\hat{F}_N^+(x,a) := \sup_{y<x} \hat{F}_N(y,a), \tag{7}$$

$$\hat{F}_N^+(x,a) := L\left[\inf_{y>x} \hat{F}_N(y,a)\right], \tag{8}$$

$$\tilde{F}_N^-(x,a) := \max(\hat{F}_N^-(x,a),0), \tag{9}$$

$$\tilde{F}_N^\pm(x,a) := \frac{1}{2}\left(\tilde{F}_N^+(x,a)+\tilde{F}_N^-(x,a)\right), \tag{10}$$

$$\hat{F}_N^\pm(x,a) = \begin{cases} \hat{F}_N^+(x,a) & \text{if } \hat{F}_N^+(x,a) \le 1/2, \\ \hat{F}_N^-(x,a) & \text{if } \hat{F}_N^-(x,a) \ge 1/2, \\ 1/2, & \text{otherwise.} \end{cases} \tag{11}$$

In what follows the index $*$ stands for any combination of symbols $\hat{}$ or $\tilde{}$, with superscripts $+$, or $-$, or $\pm$, that is, $F^*(x,a)$ can be any function of the form (5) or (7)–(11).

In the sequel we regard weighted empirical distribution functions as estimators of the distribution H_k (recall that this is the distribution of the kth component of the mixture). The process

$$B_N(x) = B_N(x,a) := \sqrt{N}\left(\hat{F}_N(x,a) - H_k(x)\right)$$

is called the empirical process for $\hat{F}_N(x,a)$, while

$$B_N^*(x) = B_N(x,a) := \sqrt{N}\left(F_N^*(x,a) - H_k(x)\right)$$

is called the empirical process for the improved weighted empirical distribution function $F_N^*(x,a)$. A point $x \in \mathbf{R}$ is called a point of growth of a distribution function H if

$$H(x+\delta) - H(x) > 0$$

for all $\delta > 0$. The set of all points of growth of H is denoted by $\operatorname{supp} H$. By $\langle a\rangle_N$ we denote the average of a row N of the matrix a. For example,

$$\langle a\rangle_N := \frac{1}{N}\sum_{j=1}^N a_{j:N}, \qquad \langle ab\rangle_N = \frac{1}{N}\sum_{j=1}^N a_{j:N} b_{j:N}.$$

We also put

$$\langle a\rangle := \lim_{N\to\infty} \langle a\rangle_N.$$

Theorem 2.1. *Let*

1) $\sup_{j,N} |a_{j:N}| < A < \infty$;
2) *the average* $\langle w^k w^m(a)^2\rangle$ *exists for all* $m = 1,\dots,M$;
3) *the function* H_m *is continuous on* $\mathbf{R}$ *for all* $m = 1,\dots,M$;
4) *for all* $m = 1,\dots,M$

$$\operatorname{supp} H_m \subseteq \operatorname{supp} H_k;$$

5) $\hat{F}_N(x,a)$ *is an unbiased estimator of* H_k, *that is,*

$$\langle aw^m\rangle_N = \mathbf{1}\{k=m\}$$

for all $m = 1,\dots,M$.

Then there are random functions $\check{B}_N(x)$ *and* $\check{B}_N^*(x)$ *such that*

1. *The distribution of* $\check{B}_N(x)$ *coincides with the distribution of* $B_N(x)$, *while the distribution of* $\check{B}_N^*(x)$ *coincides with the distribution of* $B_N^*(x)$.

2. *We have*

$$\sup_{x\in\mathbf{R}} \left|\check{B}^*_N(x) - \check{B}_N(x)\right| \to 0$$

in probability as $N \to \infty$.

Remark. To construct the processes $\check{B}_N(x)$ and $\check{B}^*_N(x)$ one needs, perhaps, to extend the main probability space.

Proof of Theorem 2.1. The result for $\hat{B}^+_N$ is proved in Theorem 2 of [3]. Since

$$\hat{F}^-_N(x,a) = L\left[1 - \hat{F}^+_{-\Xi_N}(-x)\right]$$

(by $\hat{F}^+_{-\Xi_N}(x)$ we denote the improved weighted empirical distribution function (7) constructed from the sample $-\Xi_N = (-\xi_{1:N}, \dots, -\xi_{N:N})$), Theorem 2.1 for $\hat{B}^-_N$ follows from its particular case for $\hat{B}^+_N$. Furthermore,

$$\hat{F}^-_N(x,a) \le \tilde{F}^-_N(x,a) \le \tilde{F}^\pm_N(x,a) \le \tilde{F}^+_N(x,a) \le \hat{F}^+_N(x,a).$$

Similar inequalities hold for B^*_N. This implies that Theorem 2.1 holds for $\tilde{B}^-_N$, $\tilde{B}^+_N$, and $\tilde{B}^\pm_N$. Theorem 2.1 for $\hat{B}^\pm_N$ follows from the estimates

$$\hat{F}^-_N(x,a) \le \hat{F}^\pm_N(x,a) \le \hat{F}^+_N(x,a). \qquad \square$$

In what follows we need some results on the asymptotic behavior of $B_N(x)$, $B^+_N(x)$, and $B^-_N(x)$ as $x \to \infty$ or $x \to -\infty$.

Put

$$\bar{H}(x) = \sum_{m=1}^{M} H_m(x). \tag{12}$$

Theorem 2.2. *If the assumptions of Theorem* 2.1 *are satisfied, then*

$$\sup_N \mathsf{P}\left\{\sup_{x<b} \frac{|B_N(x)|}{\bar{H}(x)^{1/2-\delta}} > \lambda\right\} \to 0 \quad \text{as } \lambda \to \infty,$$

$$\sup_N \mathsf{P}\left\{\sup_{x<b} \frac{\left|\hat{B}^+_N(x)\right|}{\bar{H}(x)^{1/2-\delta}} > \lambda\right\} \to 0 \quad \text{as } \lambda \to \infty,$$

$$\sup_N \mathsf{P}\left\{\sup_{x>b} \frac{|B_N(x)|}{\left(M - \bar{H}(x)\right)^{1/2-\delta}} > \lambda\right\} \to 0 \quad \text{as } \lambda \to \infty,$$

$$\sup_N \mathsf{P}\left\{\sup_{x>b} \frac{\left|\hat{B}^-_N(x)\right|}{\left(M - \bar{H}(x)\right)^{1/2-\delta}} > \lambda\right\} \to 0 \quad \text{as } \lambda \to \infty$$

for arbitrary b *and* δ *such that* $0 < \delta < 1/2$.

The proof of this theorem is given in Section 4.

3. The asymptotic behavior of improved weighted empirical moments

We noted above that the functions $F^*_N(x,a)$ can be represented in the following form:

$$F^*_N(x,a) = \frac{1}{N}\sum_{j=1}^{N} b^*_{j:N}\, \mathbf{1}\{\xi_{j:N} < x\}$$

where $b^*_{j:N}$ are some coefficients depending on the sample Ξ_N. Thus estimators of moments (2) constructed from $F^*(x,a)$ can be written as follows:

$$g^*_{k;N} := \int g(x)\, F^*_N(dx,a) = \frac{1}{N}\sum_{j=1}^{N} b^*_{j:N} g(\xi_{j:N}).$$

Our goal is to prove that such estimators are asymptotically normal, that is, to prove the weak convergence of $Y^*_{k;N} := \sqrt{N}(g^*_{k;N} - \bar{g})$ to the normal distribution with zero mathematical expectation.

It is shown in Theorem 2 of [2] that if the assumptions of Theorem 2.1 are satisfied and

$$\overline{g^2_m} = \int g^2(x)\, H_m(dx)$$

is finite for all H_m, $m = 1,\dots,M$, then $Y_{k;N} := \sqrt{N}(\hat{g}_{k;N} - \bar{g}) \Rightarrow Y$ where Y has the normal distribution with zero mean and variance $\sigma^2 = \langle (a)^2 d\rangle$,

$$d_{j:N} = \sum_{m=1}^{M} \overline{g^2_m} w^m_{j:N} - \left(\sum_{m=1}^{M} \overline{g_m} w^m_{j:N}\right)^2.$$

Theorem 3.1. *Let the assumptions of Theorem* 2.1 *be satisfied. If g is a function of bounded variation on* $\mathbf{R}$, *then $Y^*_{k;N} \Rightarrow Y$ for all $g^*_{k;N}$.*

Remark. It is sufficient to assume that g is a function of bounded variation on $\operatorname{supp} H_m$, $m = 1,\dots,M$.

Proof. Note that

$$Y^*_{k;N} = \sqrt{N}\left(\int g(x)\, F^*_N(dx,a) - \int g(x)\, H(dx)\right) = \int g(x)\, B^*_N(dx,a).$$

According to Theorem 2.1, there exist processes $\check{B}^*_N(x)$ and $\check{B}_N(x)$ such that the distribution of

$$\check{Y}^*_N := \int g(x)\, \check{B}^*_N(dx)$$

coincides with that of $Y^*_{k;N}$, while the distribution of $\check{Y}_N := \int g(x)\, \check{B}_N(dx)$ coincides with that of Y_N. Moreover $\sup_x \left|\check{B}^*_N(x) - \check{B}_N(x)\right| \to 0$ in probability.

Thus

$$\begin{aligned}\left|\check{Y}^*_N - \check{Y}_N\right| &= \left|\int g(x)\left(\check{B}^*_N(dx) - \check{B}_N(dx)\right)\right| = \left|\int \left(\check{B}^*_N(x) - \check{B}_N(x)\right)\, g(dx)\right| \\ &\le \operatorname{Var}_x g(x)\cdot \sup_x \left|\check{B}^*_N(x) - \check{B}_N(x)\right| \to 0\end{aligned}$$

in probability. This implies that the distribution of $\check{Y}^*_N$ (thus the distribution of $Y^*_{k;N}$) weakly converges to the same limit as $Y_{k;N}$ does, namely to Y.

The theorem is proved. □

Theorem 3.2. *Let the assumptions of Theorem* 2.1 *hold. Assume that g is a monotone continuous function and that for all $m = 1,\dots,M$ and some $0 < D, C < \infty$ and $\gamma > 0$*

$$H_m(x) \le \frac{D}{|g(x)|^{2+\gamma}} \quad \textit{for all } x < -C, \tag{13}$$

$$1 - H_m(x) \le \frac{D}{|g(x)|^{2+\gamma}} \quad \textit{for all } x > C. \tag{14}$$

Then $\hat{Y}^{\pm}_{k;N} \Rightarrow Y$ as $N \to +\infty$.

Remark. Any function g that has a bounded variation on all finite intervals can be represented as $g(x) = g^+(x) - g^-(x)$ where g^+ and g^- are monotone functions. Thus Theorem 3.2 holds for functions of bounded variation on all finite intervals if conditions (13)–(14) are satisfied for both g^+ and g^-. For example, these conditions for $g(x) = x^2$ are $H_m(x) = O(|x|^{-4-\gamma})$ as $x \to -\infty$, and $1 - H_m(x) = O(x^{-4-\gamma})$ as $x \to \infty$.

4. Proof of theorems

Lemma 4.1. *Let X be a stochastic process on the interval $[a,b]$ and let F be a nondecreasing function on the interval $[a,b]$. If*

$$\mathsf{E}\,|X(t) - X(t_1)|^\gamma |X(t_2) - X(t)|^\gamma \le (F(t_2) - F(t_1))^{2\alpha}$$

for some $\gamma > 0$ and $\alpha > 1/2$ and all $a \le t_1 < t < t_2 \le b$, then there exists a constant C depending on α and γ such that

$$\mathsf{P}\left\{\sup_{t\in[a,b]} |X(t)| \ge \varepsilon\right\} \le \mathsf{P}\{|X(a)| > \varepsilon/2\} + \mathsf{P}\{|X(b)| > \varepsilon/2\} + \frac{C}{\varepsilon^{2\gamma}}|F(b) - F(a)|^{2\alpha}$$

for all $\varepsilon > 0$.

Proof. If the assumptions of the lemma hold, then there exists a constant $K < \infty$ such that

$$\begin{aligned}(15)\qquad &\mathsf{P}\left\{\sup_{t\in[a,b]} \min\{|X(t) - X(t_1)|, |X(t_2) - X(t)|\} \ge \varepsilon\right\}\\ &\le \frac{2K}{\varepsilon^{2\gamma}}|F(b) - F(a)| \sup_{t,s\in[a,b]} |F(t) - F(s)|^{2\alpha-1} = \frac{2K}{\varepsilon^{2\gamma}}|F(b) - F(a)|^{2\alpha}\end{aligned}$$

according to inequality (15.30) in [6]. If $|X(t)| > \varepsilon$ for some point $t \in [a,b]$, then either $|X(a)| > \varepsilon/2$, or $|X(t) - X(a)| > \varepsilon/2$, or $|X(b) - X(t)| > \varepsilon/2$, or $|X(b)| > \varepsilon/2$. Thus inequality (15) implies the lemma. □

Lemma 4.2. *Let the assumptions of Theorem 2.1 hold. Then there exists $C < \infty$ independent of N and such that*

$$(16)\qquad \mathsf{P}\left\{\sup_{t<x} |B_N^+(t)| \ge \varepsilon\right\} \le \mathsf{P}\left\{\sup_{t<x} |B_N(t)| \ge \varepsilon\right\} \le C\left(\bar{H}^2(x)\varepsilon^{-4} + \bar{H}(x)\varepsilon^{-2}\right),$$

$$\mathsf{P}\left\{\sup_{t>x} |B_N^-(t)| \ge \varepsilon\right\} \le \mathsf{P}\left\{\sup_{t>x} |B_N(t)| \ge \varepsilon\right\} \le C\left((M - \bar{H}(x))^2\varepsilon^{-4} + (M - \bar{H}(x))\varepsilon^{-2}\right)$$

for all $\varepsilon > 0$ and $x \in \mathbf{R}$.

Proof. We apply Lemma 4.1 on the interval $(-\infty, x]$ for $\gamma = 2$ and $\alpha = 1$ to prove the second inequality in (16). First we estimate

$$J := \mathsf{E}(B_N(t) - B_N(t_1))^2 (B_N(t_2) - B_N(t))^2.$$

Put

$$\eta_j(x,y) := a_{j:N}\big(\mathbf{1}\{\xi_{j:N} \in [y,x)\} - \mathsf{P}\{\xi_{j:N} \in [y,x)\}\big).$$

Then

$$B_N(t) - B_N(s) = \frac{1}{\sqrt{N}} \sum_{j=1}^N \eta_j(s,t)$$

and

$$\begin{aligned}
J &= \frac{1}{N^2}\,\mathsf{E}\sum_{j,k,l,m=1}^{N}\eta_j(t,t_1)\eta_k(t,t_1)\eta_l(t_2,t)\eta_m(t_2,t)\\
&\le \frac{C}{N^2}\sum_{j\neq k}\left\{\mathsf{E}(\eta_j(t,t_1))^2(\eta_k(t_2,t))^2+\mathsf{E}\,\eta_j(t,t_1)\eta_j(t_2,t)\eta_k(t,t_1)\eta_k(t_2,t)\right\}\\
&\quad+\frac{1}{N^2}\sum_{j=1}^{N}\mathsf{E}(\eta_j(t,t_1)\eta_j(t_2,t))^2,
\end{aligned}$$

since $\mathsf{E}\,\eta_j=0$ and η_k and η_m are independent for $k\neq m$.

Since

$$\mathsf{E}\,\eta_j^2(t,t_1)\le \mathsf{E}(a_{j:N})^2\,\mathbf{1}\{\xi_{j:N}\in[t_1,t)\}\le A^2\bar H([t_1,t])\le A^2\bar H([t_1,t_2]),$$

$$\begin{aligned}
&\mathsf{E}\,\eta_j(t,t_1)\eta_j(t_2,t)=-(a_{j:N})^2\,\mathsf{P}\{\xi_{j:N}\in[t_1,t]\}\,\mathsf{P}\{\xi_{j:N}\in[t,t_2]\}\le A^2\bar H^2([t_1,t_2]),\\
&\mathsf{E}(a_{j:N})^4\left(\mathbf{1}\{\xi_{j:N}\in[t_1,t]\}-\mathsf{P}\{\xi_{j:N}\in[t_1,t]\}\right)^2\left(\mathbf{1}\{\xi_{j:N}\in[t,t_2]\}-\mathsf{P}\{\xi_{j:N}\in[t,t_2]\}\right)^2\\
&\qquad\le CA^4\bar H^3([t_1,t_2]),
\end{aligned}$$

and $\bar H(\mathbf{R})\le M$, we get $J\le C(\bar H([t_1,t_2]))^2$. Thus Lemma 4.1 implies

$$\mathsf{P}\left\{\sup_{t<x}|B_N(t)|>\varepsilon\right\}\le \mathsf{P}\left\{|B_N(x)|>\frac{\varepsilon}{2}\right\}+\frac{C}{\varepsilon^4}\left(\bar H(x)\right)^2, \tag{17}$$

since $B_N(-\infty)=\bar H(-\infty)=0$.

Then

$$\mathsf{P}\{|B_N(x)|<\varepsilon/2\}\le\frac{\mathbf{D}\,B_N(x)}{\varepsilon^2}\le\frac{A^2\bar H(x)}{\varepsilon}, \tag{18}$$

since $\mathbf{D}\,B_N(x)\le N^{-1}\sum_{j=1}^{N}(a_{j:N})^2\,\mathsf{P}\{\xi_{j:N}<x\}$. Now relations (17) and (18) imply the second inequality in (16).

The first inequality in (16) holds, since $B_N^+(x)\ge B_N(x)$ for all x and

$$\begin{aligned}
B_N^+(t)&=\sqrt{N}\sup_{y<t}\left(\hat F_N(y,a)-H_k(t)\right)=\sqrt{N}\sup_{y<t}\left(B_N(y)/\sqrt{N}+H_k(y)-H_k(t)\right)\\
&\le\sup_{y<x}B_N(y)
\end{aligned}$$

for $t<x$ in view of $H_k(y)-H_k(t)\le 0$ for $y\le t$. The other inequalities of the lemma can be proved similarly. □

Proof of Theorem 2.2. We prove the first statement of the theorem. The other statements are proved similarly. Put

$$p_\lambda:=\mathsf{P}\left\{\sup_{x<b}\frac{|B_N(x)|}{\bar H(x)^{1/2-\delta}}>\lambda\right\}$$

and let x_j be such that $\bar H(x_j)=2^{-j}$. Then

$$\mathsf{P}\left\{\sup_{x_{j+1}\le x\le x_j}\frac{|B_N(x)|}{\bar H(x)^{1/2-\delta}}>\lambda\right\}\subseteq A_j$$

where

$$A_j:=\left\{\text{for all } x<x_j,\ B_N(x)\le\lambda H^{1/2-\delta}(x_{j+1})\right\}.$$

Applying Lemma 4.2 to the event A_j for $\varepsilon_j = \lambda \bar{H}^{1/2-\delta}(x_{j+1})$ we get

$$\begin{aligned} p_\lambda &\le \sum_{j=1}^{\infty} \mathsf{P}\{A_j\} \le \sum_{j=1}^{\infty} C\left(\bar{H}^2(x_j)\varepsilon^{-4} + \bar{H}(x_j)\varepsilon^{-2}\right) \\ &= C\sum_{j=1}^{\infty}\left(\frac{2^{2j}}{\lambda^4 2^{(-2+4\delta)j}} + \frac{2^j}{\lambda^2 2^{(-1+2\delta)j}}\right) \le C(\lambda^{-4}+\lambda^{-2}) \to 0 \quad \text{as } \lambda \to +\infty. \quad \square \end{aligned}$$

Lemma 4.3. *Let the assumptions of Theorem* 2.1 *hold. If* $H_k(b) < \frac{1}{2}$ *and* $H_k(c) > \frac{1}{2}$, *then*

$$\mathsf{P}\left\{\exists x < b \colon \hat{F}_N^{\pm}(x,a) \ne \hat{F}_N^{+}(x,a)\right\} \to 0, \qquad N \to \infty,$$
$$\mathsf{P}\left\{\exists x > c \colon \hat{F}_N^{\pm}(x,a) \ne \hat{F}_N^{-}(x,a)\right\} \to 0, \qquad N \to \infty.$$

Proof. Theorem 2.1 above and Theorem 2.1 in [7] imply that

$$\sup_x \left|\hat{F}_N^{+}(x,a) - H_k(x)\right| \to 0$$

and $\sup_x\left|\hat{F}_N^{-}(x,a) - H_k(x)\right| \to 0$ in probability as $N \to \infty$. Since H_k is monotone, the assumptions of the lemma imply

$$\mathsf{P}\left\{\sup_{x<b} \hat{F}_N^{+}(x,a) > 1/2\right\} \to 0, \qquad \mathsf{P}\left\{\inf_{x>c} \hat{F}_N^{-}(x,a) < 1/2\right\} \to 0$$

as $N \to \infty$.

Taking into account (11) we complete the proof of the lemma. $\square$

Proof of Theorem 3.2. We prove the theorem for the case of an increasing function g. According to Theorem 2.1, there exist stochastic processes $\check{B}_N^{\pm}$ and $\check{B}_N$ such that $\check{B}_N^{\pm}$ has the same distribution as $\hat{B}_N^{\pm}$, and $\check{B}_N$ has the same distribution as $\hat{B}_N$. Moreover

$$\sup_x \left|\check{B}_N^{\pm}(x) - \check{B}_N(x)\right| \to 0$$

in probability as $N \to \infty$. Note that $\check{Y}_{k,N}^{\pm} = \int g(x)\, \check{B}_N^{\pm}(dx)$ has the same distribution as $\hat{Y}_{k,N}^{\pm}$, while $\check{Y}_{k,N} = \int g(x)\, \check{B}_N(dx)$ has the same distribution as $\hat{Y}_{k,N}$. Thus Theorem 3.2 follows, since $\check{Y}_{k,N}^{\pm} - \check{Y}_{k,N} \to 0$ in probability. Note that for all $b > 0$

$$J := |\check{Y}_{k,N}^{\pm} - \check{Y}_{k,N}| = \left|\int_{-\infty}^{+\infty} \left(\check{B}_N^{\pm}(x) - \check{B}_N(x)\right)\, g(dx)\right| \le J_1 + J_2 + J_3$$

where

$$J_2 = \int_{-\infty}^{-b} \left|\check{B}_N^{\pm}(x) - \check{B}_N(x)\right|\, g(dx),$$
$$J_3 = \int_{b}^{\infty} \left|\check{B}_N^{\pm}(x) - \check{B}_N(x)\right|\, g(dx).$$

Now we prove that for all $\alpha > 0$ and $\varepsilon > 0$ there are numbers b and N_0 such that

$$\sup_{N>N_0} \mathsf{P}\left\{J_2^N > \alpha\right\} < \varepsilon, \tag{19}$$

$$\sup_{N>N_0} \mathsf{P}\left\{J_3^N > \alpha\right\} < \varepsilon. \tag{20}$$

This will complete the proof of the theorem, since for all b

$$J_1^N \le (g(b) - g(-b)) \sup_{|x|<b} \left|\check{B}_N^{\pm}(x) - \check{B}_N(x)\right| \to 0$$

in probability as $N \to \infty$.

Let us prove (19). (The proof of (20) is similar.) Let b be such that

$$H_k(b) < 1/2.$$

Then Lemma 4.3 shows that it is sufficient to prove that

$$\sup_N \mathsf{P}\left\{\tilde{J}_2^N > \alpha\right\} < \varepsilon$$

for sufficiently large b where $\tilde{J}_2^N := J_{21}^N + J_{22}^N$, $J_{21}^N := \int_{-\infty}^{-b} |B_N(x)|\, g(dx)$, and

$$J_{21}^N := \int_{-\infty}^{-b} |B_N^+(x)|\, g(dx).$$

To estimate J_{21}^N we fix r and $0 < \delta < \frac{1}{2}$ such that

$$\gamma' := (2+\gamma)(1/2-\delta) > 1.$$

Theorem 2.2 implies that there exists λ such that $\sup_N \mathsf{P}\{A_N\} < \varepsilon$ for the events

$$A_N := \left\{\sup_{t<r} \frac{|B_N(t)|}{\bar{H}^{1/2-\delta}(t)} > \lambda\right\}.$$

If $\bar{A}_N$ occurs, then for all $t < r$

$$|B_N(t)| < \lambda \bar{H}^{1/2-\delta}(t) \le \frac{C}{|g(t)|^{\gamma'}}$$

and, respectively,

$$J_{21}^N \le \int_{-\infty}^{-b} \frac{C}{(g(t))^{\gamma'}}\, g(dt) < \infty$$

for $b < r$. Therefore, b can be chosen large enough that $J_{21}^N < \alpha/2$ if $\bar{A}_N$ occurs. Hence $\sup_N \mathsf{P}\{J_{21}^N > \alpha/2\} \le \sup_N \mathsf{P}\{A_N\} \le \varepsilon$.

The term J_{22}^N is estimated analogously. Therefore (19) is proved and this completes the proof of the theorem. □

Bibliography

1. R. E. Maĭboroda, *Estimators of components of a mixture with varying concentrations*, Ukrain. Mat. Zh. **48** (1996), no. 4, 562–566; English transl. in Ukrainian Math. J. **48** (1997), no. 4, 618–622. MR1417019 (97j:62055)
2. O. Kubaĭchuk, *Estimation of moments by observations from mixtures with varying concentrations*, Theory Stoch. Process. **8(24)** (2002), no. 3–4, 226–232. MR2027394 (2005g:62065)
3. R. Maĭboroda and O. Kubaĭchuk, *Asymptotic normality of improved weighted empirical distribution functions*, Teor. Imovirnost. ta Matem. Statist. **69** (2003), 89–95; English transl. in Theor. Probability and Math. Statist. **69** (2004), 95–102. MR2110908
4. V. V. Buldygin and Yu. V. Kozachenko, *Metric Characterization of Random Variables and Random Processes*, TViMS, Kiev, 1998; English transl., American Mathematical Society, Providence, RI, 2000. MR1743716 (2001g:60089)
5. A. V. Skorokhod, *Studies in the Theory of Random Processes*, Kiev University, Kiev, 1961; English transl., Addison-Wesley Publishing Co., Inc., Reading, Mass., 1965. MR0185620 (32:3082b)

6. P. Billingsley, *Convergence of Probability Measures*, Wiley, New York, 1968. MR0233396 (38:1718)
7. R. E. Maĭboroda, *Correlation analysis of mixtures. I*, Teor. Imovirnost. ta Matem. Statist. **54** (1996), 105–115; English transl. in Theor. Probability and Math. Statist. **54** (1997), 105–114. MR1644590 (99k:62102)

Department of Probability Theory and Mathematical Statistics, Mechanics and Mathematics Faculty, National Taras Shevchenko University, Academician Glushkov Avenue 6, Kyiv 03127, Ukraine

E-mail address: mre@mechmat.univ.kiev.ua

Department of Probability Theory and Mathematical Statistics, Mechanics and Mathematics Faculty, National Taras Shevchenko University, Academician Glushkov Avenue 6, Kyiv 03127, Ukraine

E-mail address: linsta@akcecc.kiev.ua

Received 22/JAN/2003

Translated by V. SEMENOV