A Weighted Weak Law of Large Numbers for Free Random Variables

Raluca Balan*
University of Ottawa

George Stoica†‡
University of New Brunswick

August 10, 2004

Abstract

We examine various conditions under which a weighted weak law of large numbers holds, in the context of noncommutative probability theory.

Keywords: weak law of large numbers, noncommutative probability theory.

1 Introduction

The weak law of large numbers (WLLN) proved by Kolmogorov is one of the most beautiful results in classical probability theory. Quite recently, this fundamental limit theorem was extended by Bercovici and Pata (1996) to the general context of noncommutative probability theory. The later context means replacing the classical probability space $(\Omega, \mathcal{F}, P)$ by a *noncommutative probability space* $(\mathcal{A}, \varphi)$, where $\mathcal{A}$ is a complex unital algebra and φ is a linear functional on $\mathcal{A}$ satisfying $\phi(1) = 1$.

The noncommutative WLLN of Bercovici and Pata was proved under the same necessary and sufficient condition as Kolmogorov's WLLN. Even more, in a paper by the same authors (see Bercovici and Pata, 1999), it is proved that if $\{X_k\}_k$ are i.i.d. random variables (in the classical sense), $\{Y_k\}_k$ are *free random variables* with the same common distribution as the X_k's, and we let

$$S_n = \frac{1}{g(n)} \sum_{k=1}^{n} X_k - M_n, \quad T_n = \frac{1}{g(n)} \sum_{k=1}^{n} Y_k - M_n$$

*Corresponding Author. *Postal address*: University of Ottawa, Department of Mathematics and Statistics, 585 King Edward Avenue, Ottawa ON, K1N 6N5, Canada. *E-mail address*: rbala348@science.uottawa.ca

†*Postal address*: University of New Brunswick, Department of Mathematical Sciences, PO Box 5050, Saint John NB, E2L 4L5, Canada. *E-mail address*: stoica@unbsj.ca

‡Both authors were supported by grants from the Natural Sciences and Engineering Research Council of Canada.

for arbitrary constants $g(n) > 0$ and M_n, then S_n converges in distribution if and only if T_n converges in distribution (but the limits may not be the same).

However, the above result does not cover the case of independent (respectively free) random variables X_k (respectively Y_k), which are *not* identically distributed, or are weighted by another constant $h(k)$.

The purpose of this paper is to generalize the WLLN of Bercovici and Pata to independent (non-identically distributed) random variables, normalized by arbitrary constants $g(n)$ and $h(k)$. As a by-product, we will obtain a necessary and sufficient condition for the weighted WLLN, similar to the one of Kolmogorov.

The paper is organized as follows. In section 2 we prove the noncommutative weighted WLLN under the same sufficient conditions as in the classical case, and we identify a necessary condition as well. In section 3 we modify slightly the result of section 2, so that the new sufficient conditions become easy to verify, under some regularity and summability assumptions on g, h. The results in section 3 are new in both classical and noncommutative settings. The appendix contains the proof of a technical lemma.

We begin by introducing the terminology and notation specific to noncommutative probability theory.

A W^**-probability space* is a pair $(\mathcal{A}, \varphi)$ be a probability space, where $\mathcal{A}$ is a complex unital von Neumann subalgebra of some $L(H)$ (the space of bounded linear operators on a Hilbert space H) and φ is a normal faithful trace. A *random variable* X is a self-adjoint operator affiliated with $\mathcal{A}$, i.e. $u(X) \in \mathcal{A}$ for any bounded Borel function u on $\mathbf{R}$. The *distribution* of a random variable X is a probability measure on $\mathbf{R}$ given by $\mu_X = \varphi \circ E_X$, where E_X is the spectral measure of X. We have

$$\varphi(u(X)) = \int_{-\infty}^{\infty} u(t) d\mu_X(t)$$

for every bounded Borel function u on $\mathbf{R}$. A sequence of random variables $\{X_n\}_{n\geq 1}$ *converges in distribution* to a probability measure ν if μ_{X_n} converges to ν weakly. We refer the reader to p. 592 (Pata, 1996a) for the definition of noncommutative *independence*, in particular *freeness*.

The analogue of the log-characteristic function associated to a probability measure μ, is the *Voiculescu transform* $\phi_\mu : \Gamma_{\alpha,\beta} \to \mathbf{C}^-$ defined implicitly by

$$G_\mu(\phi_\mu(z) + z) = \frac{1}{z}$$

where $\Gamma_{\alpha,\beta} = \{z = x + iy \in \mathbf{C}; y > \beta, |x| < \alpha y\}$ is a truncated cone with $\alpha, \beta > 0$, $\mathbf{C}^-$ is the lower half plane and $G_\mu(z) = \int_{-\infty}^{\infty} 1/(z-t) d\mu(t)$. For any constant c, we have $\phi_{\mu_{cX}}(z) = c\phi_{\mu_X}(z/c)$.

If X, Y are free random variables with distributions μ, ν respectively, then the distribution of $X+Y$ is the *free convolution* $\mu \oplus \nu$. We have $\phi_{\mu\oplus\nu} = \phi_\mu + \phi_\nu$.

For more details (e.g. continuity property), the reader is referred to Bercovici and Pata (1996), Pata (1996a).

Throughout this paper, we let $\{g(n)\}_n$ be a nondecreasing sequence of positive numbers with $g(n) \to \infty$, $\{h(n)\}_n$ a sequence of positive numbers and $f(n) = g(n)h(n)$. We denote by C a generic constant, which may be different from line to line.

2 The First Result

Our first theorem is the noncommutative analogue of the classical weighted WLLN (see Theorem 5.2.3 of Chung, 2001). Its proof is inspired by that of Theorem 2.2 (Pata, 1996a).

Theorem 2.1 *Let $\{X_k\}_{k\geq 1}$ be a sequence of independent random variables in a W^*-probability space $(\mathcal{A}, \varphi)$, and denote by μ_k the distribution of X_k. If*

$$(C1) \qquad \sum_{k=1}^{n} \mu_k(\{t : |t| > f(n)\}) \to 0$$

$$(C2) \qquad \frac{1}{g(n)^2} \sum_{k=1}^{n} \frac{1}{h(k)^2} \int_{-f(n)}^{f(n)} t^2 d\mu_k(t) \to 0$$

then there exist real constants $\{M_n\}_n$ such that the sequence

$$\frac{1}{g(n)} \sum_{k=1}^{n} \frac{X_k}{h(k)} - M_n$$

converges in distribution to the Dirac measure δ_0 at zero. Moreover, the constants M_n can be chosen to be

$$M_n = \frac{1}{g(n)} \sum_{k=1}^{n} \frac{1}{h(k)} \int_{-f(n)}^{f(n)} t \ d\mu_k(t).$$

Proof: For every $n \geq 1$, let $S_n = g(n)^{-1} \sum_{k=1}^{n} h(k)^{-1} X_k$,

$$X_{k,n}^* = X_k E_{X_k}([-f(n), f(n)]) \quad \text{and} \quad m_{k,n} = \int_{-f(n)}^{f(n)} t d\mu_k(t) \quad \text{for } k \leq n,$$

$$S_n^* = \frac{1}{g(n)} \sum_{k=1}^{n} \frac{X_{k,n}^*}{h(k)} \quad \text{and} \quad M_n = \frac{1}{g(n)} \sum_{k=1}^{n} \frac{m_{k,n}}{h(k)}.$$

We want to show that $S_n - M_n$ converges in distribution to δ_0, i.e.

$$\mu_{S_n - M_n}(\Delta_\varepsilon) \to 0 \quad \forall \varepsilon > 0 \tag{1}$$

where $\Delta_\varepsilon = \{x; |x| > \varepsilon\}$. By Proposition 3.1 (Pata, 1996a),

$$\mu_{S_n - M_n}(\Delta_\varepsilon) \leq \mu_{S_n - S_n^*}(\Delta_0) + \mu_{S_n^* - M_n}(\Delta_\varepsilon).$$

Using repeatedly Proposition 3.1 (Pata, 1996a) and condition $(C1)$, we get

$$\mu_{S_n - S_n^*}(\Delta_0) \leq \sum_{k=1}^{n} \mu_{h(k)^{-1}(X_k - X_{k,n}^*)}(\Delta_0) = \sum_{k=1}^{n} \mu_{X_k - X_{k,n}^*}(\Delta_0)$$

$$= \sum_{k=1}^{n} \mu_k(\Delta_{f(n)}) \to 0.$$

Using Chebyshev's inequality, the independence of X_k's and condition $(C2)$,

$$\mu_{S_n^* - M_n}(\Delta_\varepsilon) \leq \frac{1}{\varepsilon^2}\varphi((S_n^* - M_n)^2) \leq \frac{1}{\varepsilon^2 g(n)^2}\sum_{k=1}^{n}\frac{1}{h(k)^2}\varphi((X_{k,n})^2) =$$

$$\frac{1}{\varepsilon^2 g(n)^2}\sum_{k=1}^{n}\frac{1}{h(k)^2}\int_{-f(n)}^{f(n)} t^2 d\mu_k(t) \to 0$$

This concludes the proof of (1). $\square$

When the random variables $\{X_k\}_{k\geq 1}$ have a common distribution μ, condition $(C1)$ can be written as:

$$(C) \quad n\ \mu(\{x; |x| > f(n)\}) = o(1) \quad \text{as} \quad n \to \infty.$$

In this case, a result on p. 192 of Adler and Rosalsky (1991) says that (C) implies $(C2)$ if either one of the following sets of conditions hold:

$$(F1) \qquad f\uparrow, \quad \frac{f(n)}{n}\downarrow, \quad A(n) := \frac{1}{g(n)^2}\sum_{k=1}^{n}\frac{1}{h(k)^2} = o(1), \quad \sum_{k=1}^{n}\frac{f(k)^2}{k^2} = O(A_n^{-1})$$

$$(F2) \qquad \frac{f(n)}{n}\uparrow, \quad \sum_{k=1}^{n}\frac{1}{h(k)^2} = O\left(\frac{n}{h(n)^2}\right)$$

(Here the symbols $u_n \uparrow$ or $u_n \downarrow$ are used to indicate that the sequence $\{u_n\}_n$ is nondecreasing, respectively nonincreasing.)

Examples: 1. Let $g(n) = n^a, a \geq 0$ and $h(n) = n^b, b > 1/2$. We see that $(F1)$, respectively $(F2)$ hold, depending on whether $a + b < 1$ or $a + b \geq 1$. (In section 3, we will improve this example by requiring only $a, b \geq 0; a + b > 1/2$.)

2. If h is nonincreasing and $f(n)/n$ is nondecreasing, then $(F2)$ is satisfied. As an example we may take $g(n) = n^\rho, \rho > 1$ and $h(n) = 1/\log n$.

The next theorem says that if the random variables are free and we impose

$$(F3) \ \inf_n h(n) > 0$$

then (C) is also necessary.

Theorem 2.2 *Suppose that either* $(F1), (F3)$ *or* $(F2), (F3)$ *hold. Let* $\{X_k\}_{k\geq 1}$ *be a sequence of free identically distributed random variables with common distribution* μ*. The following are equivalent:*

(i) There exist real constants M_n *such that the sequence*

$$\frac{1}{g(n)} \sum_{k=1}^{n} \frac{X_k}{h(k)} - M_n$$

converges in distribution to the Dirac measure δ_0 *at zero.*

(ii) The measure μ *satisfies* (C)*.*

Moreover, if (ii) is satisfied the constants M_n *in (i) can be chosen as in Theorem 2.1 (with* $\mu_k = \mu$*).*

Proof: It remains to prove that $(i) \Rightarrow (ii)$.

For every $n \geq 1$, let ν_n be the distribution of $\sum_{k=1}^n X_k/d_{k,n} - M_n$, where $d_{k,n} = g(n)h(k)$ for $k \leq n$. By Proposition 1 (Bercovici and Pata, 1996), (i) implies that

$$\lim_{y\to\infty} \frac{\phi_{\nu_n}(iy)}{y} = 0 \quad \text{uniformly in } n. \tag{2}$$

Note that $\phi_{\nu_n}(iy) = \sum_{k=1}^n \phi_\mu(iyd_{k,n})/d_{k,n} - M_n$ and

$$\frac{\Im\phi_{\nu_n}(iy)}{y} = \sum_{k=1}^{n} \frac{\Im\phi_\mu(iyd_{k,n})}{yd_{k,n}}. \tag{3}$$

Using Proposition 2.5 (Bercovici and Pata, 1999), we have

$$\phi_\mu(z) = z^2 \left[G_\mu(z) - \frac{1}{z}\right](1 + v(z))$$

where $v(z) \to 0$ as $|z| \to \infty$ nontangentially. Since $\phi_\mu : \Gamma_{\alpha,\beta} \to \mathbf{C}^-$, we have

$$\Im\phi_\mu(iy) = -y^2 \Im\left[G_\mu(iy) - \frac{1}{iy}\right](1 + v(iy)) \leq 0, \quad \forall y \geq \beta \tag{4}$$

which implies $\Im[G_\mu(iy) - 1/(iy)](1 + v(iy)) \geq 0$ for all $y \geq \beta$.

Since $g(n) \to \infty$ and $(F3)$ holds, there exists N such that $d_{k,n} \geq mg(n) \geq \beta$ for all $n \geq N, k \leq n$. From (3) and(4), we get

$$\begin{aligned} \frac{\Im \phi_\mu(iy)}{y} &= -y \sum_{k=1}^{n} d_{k,n} \Im \left[G_\mu(iyd_{k,n}) - \frac{1}{iyd_{k,n}} \right] (1 + v(iyd_{k,n})) \\ &\leq -yf(n) \Im \left[G_\mu(iyf(n)) - \frac{1}{iyf(n)} \right] (1 + v(iyf(n))) \leq 0. \end{aligned}$$

From (2), we get

$$\lim_{y \to \infty} yf(n) \Im \left[G_\mu(iyf(n)) - \frac{1}{iyf(n)} \right] (1 + v(iyf(n))) = 0 \quad \text{uniformly in } n.$$

Since $f(n) \geq mg(1)$ for all n and $|v(iyf(n))| \leq 1/2$ for n, y large enough, we conclude that

$$\lim_{y \to \infty} \Im \left[G_\mu(iyf(n)) - \frac{1}{iyf(n)} \right] = 0 \quad \text{uniformly in } n$$

which implies

$$\lim_{y \to \infty} \Im \left[G_\mu(iy) - \frac{1}{iy} \right] = 0.$$

The conclusion follows immediately since

$$\Im \left[G_\mu(iy) - \frac{1}{iy} \right] = \int_{-\infty}^{\infty} \left(\frac{-y}{t^2 + y^2} + \frac{1}{y} \right) d\mu(t) \geq \frac{1}{2} y \ \mu(\{t; |t| \geq f(y)\}).$$

□

3 The Second Result

The following result is obtained by applying Theorem 2.1 to the random variables $\tilde{X}_k = X_k / h(k)$ and the sequences $\tilde{g}(n) = g(n)$, $\tilde{h}(n) = 1$. Note that in this case $\tilde{X}_k$ has distribution $\tilde{\mu}_k$ defined by $\tilde{\mu}_k(B) := \mu_k(h(k)B)$.

Theorem 3.1 *Let $\{X_k\}_{k \geq 1}$ be a sequence of independent random variables in a W^*-probability space, and denote with μ_k the distribution of X_k. If*

$$(\widetilde{C}1) \qquad \sum_{k=1}^{n} \mu_k(\{t; |t| > g(n)h(k)\}) \to 0$$

$$(\widetilde{C}2) \qquad \frac{1}{g(n)^2} \sum_{k=1}^{n} \frac{1}{h(k)^2} \int_{-g(n)h(k)}^{g(n)h(k)} t^2 d\mu_k(t) \to 0$$

then there exist real constants $\{\tilde{M}_n\}_n$ such that the sequence

$$\frac{1}{g(n)}\sum_{k=1}^{n}\frac{1}{h(k)}X_k - \tilde{M}_n$$

converges in distribution to the Dirac measure δ_0 at zero. Moreover, the constants $\tilde{M}_n$ can be chosen to be

$$\tilde{M}_n = \frac{1}{g(n)}\sum_{k=1}^{n}\frac{1}{h(k)}\int_{-g(n)h(k)}^{g(n)h(k)} t \; d\mu_k(t).$$

In what follows we will suppose that X_k's have common distribution μ, g can be extended to a positive nondecreasing function on $(0,\infty)$ and h can be extended to a positive function on $(0,\infty)$. Let $f(x) := g(x)h(x)$.

The following assumptions on g, h will be used, in addition to $(F3)$:

$(F4)$ the inverse f^{-1} of f exists and satisfies $\lim_{t\to\infty} f^{-1}(t) = \infty$

$(F5)$ $\displaystyle\sum_{k=1}^{n}\frac{1}{f^{-1}(g(n)h(k))} = O(1)$

$(F6)$ f is regularly varying at ∞ with index $\rho > 1/2$, i.e. for every $\lambda > 0$ $\lim_{x\to\infty} f(\lambda x)/f(x) = \lambda^\rho$

Examples: 1. Let $g(x) = x^a, h(x) = x^b$ with $a, b \geq 0, a+b > 1/2$. In this case $f^{-1}(x) = x^{1/(a+b)}$ and $(F3)$-$(F6)$ hold: in the case of $(F5)$ we have

$$\sum_{k=1}^{n}(n^a k^b)^{-1/(a+b)} = n^{-a/(a+b)}\sum_{k=1}^{n}k^{-b/(a+b)} \leq n^{-a/(a+b)}\cdot Cn^{1-b/(a+b)} = C.$$

2. If h is nonincreasing and f is nondecreasing, then $(F5)$ holds:

$$\sum_{k=1}^{n}\frac{1}{f^{-1}(g(n)h(k))} \leq \sum_{k=1}^{n}\frac{1}{f^{-1}(f(n))} = \sum_{k=1}^{n}\frac{1}{n} = 1.$$

As an example we may take $g(x) = x^\rho, \rho > 1$ and $h(x) = (x+1)/x$. In this case $f(x) = x^{\rho-1}(x+1)$ is strictly increasing and $(F3)$-$(F6)$ hold.

We consider also a stronger form of condition (C):

$$(\widetilde{C})\ \tau(t) := f^{-1}(t)\ \mu(\{x; |x| > t\}) = o(1) \quad \text{as} \quad t\to\infty.$$

Proposition 3.2 *Under $(F3)$-$(F5)$, $(\widetilde{C})$ implies $(\widetilde{C}1)$.*

Proof: Let $d_{k,n} := g(n)h(k)$ for $k \leq n$. Let $\varepsilon > 0$ be arbitrary. By $(\widetilde{C})$, there exists $T = T_\varepsilon$ such that $\tau(t) < \varepsilon$ for all $t \geq T$. Since $g(n) \to \infty$ and $(F3)$ holds, there exists $N = N_\varepsilon$ such that $d_{k,n} \geq mg(n) \geq T$ for all $n \geq N, k \leq n$. Therefore, for every $n \geq N$

$$\sum_{k=1}^{n} \mu(\{x; |x| > d_{k,n}\}) = \sum_{k=1}^{n} \frac{1}{f^{-1}(d_{k,n})} \tau(d_{k,n}) \leq \varepsilon \sum_{k=1}^{n} \frac{1}{f^{-1}(d_{k,n})} \leq C\varepsilon,$$

where we used $(F5)$ for the last inequality. This concludes the proof. $\square$

The following lemma generalizes Lemma 4.3.(iii) (Bercovici and Pata, 1996) to the case of invertible, regularly varying functions f. Its proof is given in the appendix.

Lemma 3.3 *Under $(F4)$ and $(F6)$, if $(\widetilde{C})$ holds, then for every $k \geq 2$,*

$$v_k(y) := \frac{f^{-1}(y)}{y^k} \int_{-y}^{y} |t|^k d\mu(t) = o(1) \quad \text{as} \quad y \to \infty.$$

Using this lemma, we obtain the following result.

Proposition 3.4 *Under $(F3)$-$(F6)$, $(\widetilde{C})$ implies $(\widetilde{C}2)$.*

Proof: Let $d_{k,n} := g(n)h(k)$ for $k \leq n$. Let $\varepsilon > 0$ be arbitrary. By Lemma 3.3, there exists $Y = Y_\varepsilon$ such that $v_2(y) \leq \varepsilon, \forall y \geq Y$. Since $g(n) \to \infty$ and $(F3)$ holds, there exists $N = N_\epsilon$ such that $d_{k,n} > Y$ for all $n \geq N, k \leq n$. Hence, using $(F5)$ we have

$$\sum_{k=1}^{n} \frac{1}{d_{k,n}^2} \int_{-d_{k,n}}^{d_{k,n}} t^2 d\mu(t) = \sum_{k=1}^{n} \frac{v_2(d_{k,n})}{f^{-1}(d_{k,n})} \leq \varepsilon \sum_{k=1}^{n} \frac{1}{f^{-1}(d_{k,n})} \leq C\varepsilon$$

for all $n \geq N$, which concludes the proof. $\square$

The next theorem says that under $(F3) - (F6)$, $(\widetilde{C})$ is a necessary and sufficient condition for the weighted WLLN. Sufficiency is obtained immediately from Theorem 3.1 and Propositions 3.2, 3.4; necessity follows exactly as in Theorem 2.2.

Theorem 3.5 *Suppose that $(F3)$-$(F6)$ hold. Let $\{X_k\}_{k\geq 1}$ be a sequence of free identically distributed random variables with common distribution μ. The following are equivalent:*

(i) There exist real constants $\{\tilde{M}_n\}_n$ such that the sequence

$$\frac{1}{g(n)} \sum_{k=1}^{n} \frac{1}{h(k)} X_k - \tilde{M}_n$$

converges in distribution to the Dirac measure δ_0 at zero.

(ii) The measure μ satisfies $(\widetilde{C})$.

Moreover, if (ii) is satisfied the constants $\tilde{M}_n$ in (i) can be chosen as in Theorem 3.1 (with $\mu_k = \mu$).

Concluding Remarks: (a) Theorems 2.2 and 3.5 extend several results in noncommutative probability theory, such as Kolmogorov WLLN (cf. Bercovici and Pata, 1996) and Marcinkiewicz WLLN (cf. Pata, 1996a), and in addition give new such WLLN's considering for instance regularly varying weights.

(b) Propositions 3.2, 3.4 show that under $(F3)-(F6)$, $(\widetilde{C})$ is also a sufficient condition for the weighted WLLN in the classical sense, which seems to be a new result. See Theorem 1.3 (Gut, 2004) which treats the case $h(n) = 1$.

(c) In view of the the central limit theorem for free random variables (cf. Pata, 1996b), the index ρ in $(F6)$ has to be strictly larger than $1/2$.

(d) The case of logarithmic averages (i.e. $g(n) = \log n, h(n) = n$) is not covered either by conditions $(F1)-(F2)$, nor by conditions $(F3)-(F6)$. In the classical theory, condition

$$\int_{-\infty}^{\infty} f^{-1}(t) d\mu(t) < \infty, \quad f(t) := t \log t$$

is known to be necessary and sufficient for the *strong* LLN (see Jaite, 2004). However, even in this setting, it is not clear whether the WLLN for logarithmic averages holds, under $(\widetilde{C})$ alone.

A Appendix

Proof of Lemma 3.3: Using integration by parts, we have

$$\int_{-y}^{y} |t|^k d\mu(t) = -y^k \mu(\{x; |x| > y\}) + k \int_0^y t^{k-1} \mu(\{x; |x| > t\}) dt.$$

The result will follow once we prove that

$$\frac{f^{-1}(y)}{y^k} \int_0^y t^{k-1} \mu(x; |x| > t) dt = o(1) \quad \text{as} \quad y \to \infty. \tag{5}$$

Note that f^{-1} is regularly varying at ∞ with index $1/\rho$. Using Karamata's Representation Theorem (see Theorems 1.3.1, 1.4.1, Bingham, Goldie and Teugels, 1987), respectively Potter's Theorem (see Theorem 1.5.6.(iii), Bingham, Goldie and Teugels, 1987), we know that for every $\delta > 0, A > 1$ there exists $C = C_\delta > 0, Y = Y_{\delta,A} > 0$ such that

$$f^{-1}(y) \le C y^{(1/\rho)+\delta}, \quad \forall y \ge Y \tag{6}$$

$$\frac{f^{-1}(y)}{f^{-1}(t)} \le A\left(\frac{y}{t}\right)^{(1/\rho)+\delta}, \forall y \ge t \ge Y \tag{7}$$

Let $0 < \delta < k - (1/\rho)$ and $A > 1$ be fixed. Let $\varepsilon > 0$ be arbitrary. By $(\widetilde{C})$, there exists $N = N_\varepsilon > Y$ such that $\tau(t) < \varepsilon$, $\forall t > N$. Using (6),

$$\frac{f^{-1}(y)}{y^k}\int_0^N t^{k-1}\mu(x;|x|>t)dt \le N^{k-1}\frac{f^{-1}(y)}{y^k} \le N^{k-1}Cy^{(1/\rho)+\delta-k} \le \varepsilon \tag{8}$$

for y large enough. Using (7),

$$\frac{f^{-1}(y)}{y^k}\int_N^y t^{k-1}\mu(x;|x|>t)dt \le \varepsilon\frac{f^{-1}(y)}{y^k}\int_N^y \frac{t^{k-1}}{f^{-1}(t)}dt \le \frac{A\epsilon}{y^k}\int_N^y t^{k-1}\left(\frac{y}{t}\right)^{(1/\rho)+\delta}dt$$

$$= \frac{A\varepsilon}{y^\alpha}\int_N^y t^{\alpha-1}dt = \frac{A\varepsilon}{\alpha}\left[1-\left(\frac{N}{y}\right)^\alpha\right] \le \frac{A\varepsilon}{\alpha} \tag{9}$$

where $\alpha := k - (1/\rho) - \delta > 0$. The proof of (5) is complete by (8)- (9). □

References

[1] Adler, A. and Rosalsky, A. (1991), On the weak law of large numbers for normed weighted sums of i.i.d. random variables, *Intern. J. Math. & Math. Sci.* **14**, 191-202.

[2] Bingham, N. H., Goldie, C. M. and Teugels, J. L. (1987), *Regular Variation* (Cambridge University Press, Cambridge).

[3] Bercovici, H. and Pata, V. (1996), The law of large numbers for free identically distributed random variables, *Ann. Probab.* **24**, 453-465.

[4] Bercovici, H. and Pata, V. (1999), Stable laws and domains of attraction in free probability theory (with an appendix by Philippe Biane). *Ann. Math.* **149**, 1023-1060.

[5] Chung, K. L. (2001), *A Course in Probability Theory*, Third Edition (Academic Press, San Diego).

[6] Gut, A. (2004), An extension of Feller's weak law of large numbers. Preprint.

[7] Jaite, R. (2003), On the strong law of large numbers, *Ann. Probab.* **31**, 409-412.

[8] Pata, V. (1996a), A generalized weak law of large numbers for noncommuting random variables, *Indiana Univ. Math. J.* **45**, 591-601.

[9] Pata, V. (1996b), The central limit theorem for free additive convolution, *J. Funct. Anal.* **140**, 359-380.