

## ON QUOTIENTS OF MOVING AVERAGE PROCESSES WITH INFINITE MEAN

MAREK KANTER<sup>1</sup>

**ABSTRACT.** In this paper it is shown that one can estimate the sum of the weights used to form a stationary moving average stochastic process based on nonnegative random variables by taking the limit in probability of suitable quotients, even when the random variables involved have infinite expectation.

**1. Introduction.** Suppose we are given a two-sided infinite sequence  $X = (X_k | k = 0, \pm 1, \pm 2, \dots)$  of independent, identically distributed, nonnegative random variables such that  $E(X_k) = +\infty$  for all  $k$ . For certain infinite sequences of constants  $a = (a_k | k = 0, \pm 1, \dots)$  it is possible to form the moving average process

$$Y_n = \sum_k a_{n-k} X_k$$

which clearly is a stationary process. It is known also to be an ergodic process (see [2, p. 158]). However, the study of the ergodic properties of the process  $Y$  is made difficult by the fact that  $E(|Y_n|) = \infty$ . For instance the usual ergodic theorem cannot be applied to study the behaviour of the normed sums  $N^{-1} \sum_{n=1}^N Y_n$ .

Suppose now that  $b = (b_k | k = 0, \pm 1, \pm 2, \dots)$  is another sequence of constants such that it is possible to form the moving average process

$$Z_n = \sum_k b_{n-k} X_k.$$

The ratios  $\sum_{n=1}^N Y_n / \sum_{n=1}^N Z_n$  have also been studied, but again with the restriction that means are finite. We show in this paper that when considering ratios as above, the restriction that means are finite can be lifted, and that by fairly elementary arguments one can in fact prove

$$\lim_{N \rightarrow \infty} \frac{\sum_{n=1}^N Y_n}{\sum_{n=1}^N Z_n} = \frac{\Sigma(a)}{\Sigma(b)}$$

where the limit is taken in probability. (For any sequence  $a = (a_k | k = 0, \pm 1, \pm 2, \dots)$  we let  $\Sigma(a) = \sum_k a_k$ . In order that the moving average processes  $Y_n$  and  $Z_n$  exist it turns out that the sums  $\Sigma|a_k|$  and  $\Sigma|b_k|$  are finite.) The above

---

Received by the editors December 30, 1974.

*AMS (MOS) subject classifications* (1970). Primary 60G10, 60G50; Secondary 46E30.

*Key words and phrases.* Moving average process, stationary process, ergodic theorems, quotients of independent random variables, metric linear space.

<sup>1</sup>Research partially supported by NRC Grant No. A8753 and Quebec Action Concertee.

© American Mathematical Society 1976

statement, of course, is only asserted for  $\Sigma(b) \neq 0$ .

We will show that, unlike the usual ergodic theorems for stationary integrable random variables, our result cannot be refined to provide a.s. convergence.

Our work on this paper has been motivated by a previous study of autoregressive stationary processes with infinite variance. (See [5].)

We end this introduction with some remarks on notation.

$l$  stands for the set of all two-sided sequences  $a = (a_k | k = 0, \pm 1, \pm 2)$  of real numbers.

$l^+ = \{a | a \in l; a_k \geq 0 \text{ for all } k\}$ .

$l_0 = \{a | a \in l; a_k = 0 \text{ for } |k| \text{ sufficiently large}\}$ .

$l_0^+ = l_0 \cap l^+$ .

$l^\alpha = \{a | a \in l; \sum_k |a_k|^\alpha < \infty\}$ .

For  $a \in l^\alpha$  we define  $\rho_\alpha(a) = \sum_k |a_k|^\alpha$ . We note that  $\rho_\alpha(a + b) \leq \rho_\alpha(a) + \rho_\alpha(b)$  if  $\alpha \in (0, 1]$ .

$\delta_N$  stands for that sequence  $a$  in  $l_0^+$  whose coordinates are all zero except for  $k = 1, \dots, N$  where  $a_k = 1$ .  $\hat{\delta}_N$  stands for that sequence  $a$  in  $l_0^+$  whose coordinates are all zero except for  $k = -N, \dots, -1$  where  $a_k = 1$ .

If  $a \in l$  we define  $a^+$  to be that sequence  $c \in l$  such that  $c_k = a_k$  if  $a_k \geq 0$  and  $c_k = 0$  if  $a_k < 0$ . Similarly define  $a^-$  to be that sequence  $d \in l$  such that  $d_k = -a_k$  if  $a_k \leq 0$  and  $d_k = 0$  if  $a_k > 0$ . We note that  $a = a^+ - a^-$  and that  $\rho_\alpha(a) = \rho_\alpha(a^+) + \rho_\alpha(a^-)$ .

If  $a, b \in l^1$  we define  $a * b \in l^1$  to be that sequence  $c$  in  $l^1$  whose  $k$ th term  $c_k$  is equal to  $\sum_j a_{k-j} b_j$ .

Finally, if  $a \in l$  and  $X = (X_k | k = 0, \pm 1, \pm 2, \dots)$  are random variables such that  $\sum_k a_k X_k$  converges in probability, we will write  $a \circ X = \sum_k a_k X_k$ . As an example of our notation we rewrite  $\sum_{n=1}^N (\sum_k a_{n-k} X_k)$  as  $(\hat{\delta}_N * a) \circ X$ .

We can also write  $\sum_{n=-N}^N (\sum_k a_{n-k} X_k)$  as  $(\delta_N * a) \circ X$ . To spare the notation we will state our theorems in terms of the random variables  $(\delta_N * a) \circ X$  though they can just as easily be stated and proved for the random variables  $(\hat{\delta}_N * a) \circ X$ .

**2. Metric linear spaces associated with positive random variables.** Let  $X = (X_k | k = 0, \pm 1, \pm 2, \dots)$  be a two-sided infinite sequence of positive, independent and identically distributed random variables. We define the function  $p_X(s)$  on  $[0, \infty)$  by

$$(2.1) \quad E(e^{-sX_0}) = \exp(-p_X(s)).$$

We wish to define the random variable  $a \circ X$  for vectors  $a \in l$ . This cannot be done for all  $a \in l$ , so let us restrict our attention to those  $a \in l$  for which  $\rho_X(a) < \infty$ , where  $\rho_X(a) = \sum_k p_X(|a_k|)$ . (We shall denote this subset of  $l$  by  $l^X$ .)

It is clear from (2.1) that if  $a \in l^+ \cap l^X$  then  $a \circ X$  can be defined with

$$(2.2) \quad E(e^{-(a \circ X)}) = \exp(-\rho_X(a)).$$

For general  $a \in l^X$ , we write  $a = a^+ - a^-$  as before and define  $a \circ X = a^+ \circ X - a^- \circ X$ . (Note that  $a^+$  and  $a^-$  are in  $l^X$  since  $p_X(\cdot)$  is a nondecreasing function on  $[0, \infty)$ .)

It is useful to keep the special case of positive stable random variables of

index  $\alpha \in (0, 1)$  in mind throughout this discussion. In that case  $p_X(s) = Ks^\alpha$  for  $s \in [0, \infty)$  where  $K$  is a positive constant (which we assume to be 1 for convenience). Similarly,  $\rho_X$  turns out to be  $\rho_\alpha$  and  $l^X$  is evidently  $l^\alpha$ . (See [3] for background material on positive stable random variables.)

LEMMA 2.1.  $(l^X, \rho_X)$  is a metric linear space. Furthermore for  $a, b \in l^X$  we have  $|\rho_X(a^+) - \rho_X(b^+)| \leq \rho_X(b - a)$ .

PROOF. It is a well-known fact that  $p_X$  is concave on  $[0, \infty)$  (see, e.g., [6, p. 54]). We conclude from this that  $p_X$  is subadditive on  $[0, \infty)$  by arguments such as in [4, p. 41]. It is in turn clear that  $\rho_X$  is a metric and  $l^X$  is a linear space.

The second assertion follows directly from the inequality  $|\rho_X(u^+) - \rho_X(v^+)| \leq p_X(|u - v|)$  valid for  $u, v \in R$  and which follows directly from the fact that  $p_X$  is nondecreasing and subadditive on  $(0, \infty)$ . Q.E.D.

LEMMA 2.2.  $l^X \subset l^1$ .

PROOF. If  $a \in l^+$  and  $a \circ X$  exists then  $a \in l^1$  by an easy application of the 3 series theorem [8, p. 237]. The rest is obvious. Q.E.D.

We now list some facts which show how the metric  $\rho_X$  on  $l^X$  interrelates with the random variables  $a \circ X$ . We note that a sequence of random variables  $V_n$  converges to 0 in probability (written  $V_n \xrightarrow{P} 0$ ) if and only if  $E(|V_n|/1 + |V_n|) \rightarrow 0$ . We also note that  $V_n$  converges to  $+\infty$  in probability (written  $V_n \xrightarrow{P} +\infty$ ) if and only if for all  $\varepsilon > 0$  and  $x > 0$ ,  $P[V_n < x] < \varepsilon$  for  $n$  sufficiently large.

FACT 2.1. For any  $\varepsilon > 0$ ,  $\exists \delta > 0$  such that  $\rho_X(a) < \delta$  implies

$$E|a \circ X| / (1 + |a \circ X|) < \varepsilon \quad \text{for all } a \in l^X.$$

For any  $\varepsilon > 0$  and  $x > 0$  there exists an  $M > 0$  such that for  $a \in l^X \cap l^+$  we have  $\rho_X(a) > M \Rightarrow P[a \circ X < x] < \varepsilon$ .

Fact 2.1 follows directly from (2.2). Fact 2.2 follows from the following lemma which is of interest in itself.

LEMMA 2.3. Let  $V_n$  and  $W_n$  be a sequence of nonnegative random variables with  $V_n$  independent of  $W_n$  for all  $n$  and  $W_n > 0$  a.s. Let

$$E(e^{-sV_n}) = \exp(-p_{V_n}(s)) \quad \text{and} \quad E(e^{-sW_n}) = \exp(-p_{W_n}(s))$$

for all  $s \in (0, \infty)$ . Suppose  $\lim_{n \rightarrow \infty} (p_{W_n}(s)/p_{V_n}(s))$  is equal to 0 for all  $s \in (0, \infty)$ . Then  $(V_n/W_n) \xrightarrow{P} +\infty$ .

PROOF. For  $v \geq 0$  and  $w > 0$  we can write

$$(1 + v/w)^{-1} = \int_0^\infty w e^{-wr} e^{-vr} dr.$$

It follows immediately that

$$E((1 + (V_n/W_n))^{-1}) = \int_0^\infty E(W_n e^{-W_n r}) E(e^{-V_n r}) dr.$$

It is trivial to check that  $p_{W_n}(r)$  is differentiable for all  $r > 0$  and that

$$E(W_n e^{-W_n r}) = p'_{W_n}(r) \exp(-p_{W_n}(r));$$

hence we can write

$$(2.3) E((1 + (V_n/W_n))^{-1}) = \int_0^\infty p'_{W_n}(r) \exp(-p_{W_n}(r)) \exp(-p_{V_n}(r)) dr.$$

Let us define the function  $K_n: (0, \infty) \rightarrow (0, \infty)$  by

$$K_n(p_{V_n}(r)) = p_{V_n}(r)/p_{W_n}(r).$$

By the hypotheses of the theorem we know that  $K_n(s) \rightarrow +\infty$  for all  $s \in (0, \infty)$ . By a simple change of variables we rewrite (2.3) as

$$E((1 + (V_n/W_n))^{-1}) = \int_0^\infty e^{-s(1+K_n(s))} ds$$

and immediately conclude (by the Lebesgue dominated convergence theorem) that  $E((1 + (V_n/W_n))^{-1}) \rightarrow 0$ . This implies that  $V_n/W_n \xrightarrow{P} +\infty$ . Q.E.D.

To obtain Fact 2.2 from Lemma 2.3 we simply note that if  $V_n$  and  $W_n$  are positive random variables with  $V_n/W_n \xrightarrow{P} +\infty$  and  $V_n \xrightarrow{P} +\infty$ , then  $V_n - W_n \xrightarrow{P} +\infty$ .

The following sequence of lemmas interrelate the metric  $\rho_X$  on  $l^X$  with the convolution structure of  $l^X$ . We note first the trivial fact that for  $a \in l^X$  we have  $\rho_X(\delta_N * a) \leq \rho_X(\delta_N)\rho_X(a)$  which follows immediately from the subadditivity of  $\rho_X$ . (In the case of positive stable random variables we have the more general inequality  $\rho_\alpha(a * b) \leq \rho_\alpha(a)\rho_\alpha(b)$  for  $0 < \alpha < 1$  and  $a, b \in l^\alpha$ .)

From now on we assume that our random variables  $X = (X_k | k = 0, \pm 1, \dots)$  have been renormalized by division with a positive constant so as to make  $p_X(1) = 1$  and hence  $\rho_X(\delta_N) = N$ .

LEMMA 2.4. Let  $a \in l^X$  with  $\Sigma(a) > 0$ . Then

$$\lim_{N \rightarrow \infty} N^{-1} \rho_X((\delta_N * a)^+) = p_X(\Sigma(a)).$$

PROOF. The lemma is easily verified for sequences  $a \in l_0$ . Now for an arbitrary element  $a \in l^X$  and for  $\varepsilon > 0$  let  $a_\varepsilon \in l_0$  have the property that  $\Sigma(a) = \Sigma(a_\varepsilon)$  and  $\rho_X(a - a_\varepsilon) \leq \varepsilon$ . We now compute

$$(2.4) \quad \begin{aligned} & N^{-1} |\rho_X((\delta_N * a)^+) - N p_X(\Sigma(a))| \\ & \leq N^{-1} |\rho_X((\delta_N * a)^+) - \rho_X((\delta_N * a_\varepsilon)^+)| \\ & \quad + N^{-1} |\rho_X((\delta_N * a_\varepsilon)^+) - N p_X(\Sigma(a_\varepsilon))|. \end{aligned}$$

By Lemma 2.1 we can bound the first term on the right-hand side of (2.4) by  $N^{-1} \rho_X(\delta_N * (a - a_\varepsilon))$  which is less than or equal to  $N^{-1} \rho_X(\delta_N) \rho_X(a - a_\varepsilon)$ . Since  $\rho_X(\delta_N) = N$  we conclude that the first term on the right-hand side of (2.4) is dominated by  $\varepsilon$ . If we now take  $\limsup_{N \rightarrow \infty}$  of both sides of (2.4), then (noting that

$$\limsup_{N \rightarrow \infty} N^{-1} |\rho_X((\delta_N * a_\varepsilon)^+) - N p_X(\Sigma(a_\varepsilon))| = 0$$

since  $a_\varepsilon \in l_0$ ) we conclude that

$$\limsup_{N \rightarrow \infty} N^{-1} |\rho_X((\delta_N * a)^+) - N p_X(\Sigma(a))| \leq \varepsilon.$$

Since  $\varepsilon > 0$  was arbitrary, we get the required conclusion. Q.E.D.

The same reasoning as in Lemma 2.4 demonstrates the following fact:

LEMMA 2.5. *Let  $a \in l^X$  with  $\Sigma(a) > 0$ . Then*

$$\lim_{N \rightarrow \infty} N^{-1} \rho_X(\delta_N * a) = \rho_X\left(\sum(a)\right).$$

COROLLARY 2.1. *Let  $a \in l^X$  with  $\Sigma(a) > 0$ . Then*

$$\lim_{N \rightarrow \infty} \rho_X((\delta_N * a)^+) / \rho_X((\delta_N * a)^-) = +\infty.$$

PROOF. Just remember that

$$\rho_X((\delta_N * a)) = \rho_X((\delta_N * a)^+) + \rho_X((\delta_N * a)^-)$$

and apply Lemmas 2.4 and 2.5. Q.E.D.

### 3. The main theorem.

THEOREM 3.1. *Let  $X = (X_k | k = 0, \pm 1, \dots)$  be a two-sided sequence of positive, identically distributed, independent random variables. Let  $a, b \in l^X$  with  $\Sigma(b) \neq 0$ . Then*

$$(3.1) \quad \lim_{N \rightarrow \infty} \frac{(\delta_N * a) \circ X}{(\delta_N * b) \circ X} = \frac{\Sigma(a)}{\Sigma(b)}$$

where the limit is taken in probability.

PROOF. We first demonstrate the theorem in the special case when all the coordinates  $b_n$  of  $b = (b_n)$  are nonnegative. In that case for any real  $s$  we can write

$$(3.2) \quad P\left[\frac{(\delta_N * a) \circ X}{(\delta_N * b) \circ X} > s\right] = P[(\delta_N * (a - sb)) \circ X > 0].$$

We shall first show that if  $s > \Sigma(a)/\Sigma(b)$  then the right-hand side of (3.2) tends to 0 with  $N$ .

In fact,  $\Sigma(a - sb) < 0$ , hence  $\rho_X((\delta_N * (a - sb))^-) \rightarrow +\infty$  by Lemma 2.4 applied to  $(sb - a)$ ; also  $\rho_X((\delta_N * (a - sb))^-) / \rho_X((\delta_N * (a - sb))^+) \rightarrow +\infty$  with  $N$  by Corollary 2.1 applied to  $sb - a$ . We conclude that  $(\delta_N * (a - sb)) \circ X$  tends to  $-\infty$  in probability by Fact 2.2 applied to  $(sb - a)$ , and hence that the left-hand side of (3.2) tends to 0 with  $N$ .

In a similar way, we can show that if  $s < \Sigma(a)/\Sigma(b)$  then  $P[(\delta_N * a) \circ X / ((\delta_N * b) \circ X) < s]$  tends to 0 as  $N \rightarrow \infty$ . We conclude that

$$\lim_{N \rightarrow \infty} (\delta_N * a) \circ X / (\delta_N * b) \circ X = \Sigma(a) / \Sigma(b)$$

in the special case when all the coordinates of  $b$  are nonnegative.

To prove the theorem in full generality we invert our results so far; i.e. we know that for  $\Sigma(a) \neq 0$  and  $b$  with all coordinates nonnegative we have

$$\lim_{N \rightarrow \infty} (\delta_N * b) \circ X / (\delta_N * a) \circ X = \Sigma(b) / \Sigma(a)$$

Both sides of the last equation are linear in  $b$ ; hence the restriction on  $b$  can be removed by linearity. The general case now follows by reinverting with the

exception of the special case that  $\Sigma(a) = 0$ . This special case is taken care of by applying (3.1) successively to the positive part of  $a$ , and then to the negative part of  $a$ , and then using the fact that (3.1) is linear in  $a$ . Q.E.D.

REMARK. Theorem 3.1 is true if  $X_i$  are only assumed nonnegative (and not identically 0). This follows since  $P[(\delta_N * c) \circ X = 0]$  evidently tends to 0 for  $c \in l^X \cap l^+$ ,  $c \neq 0$ . The details are straightforward.

**4. On almost sure convergence.** In this section we will show the impossibility of proving almost sure convergence in Theorem 3.1, in the special case when the random variables  $X_k$  are all positive stable of index  $\alpha$  in  $(0, 1)$ . By looking at the special case when  $a$  is  $\delta_1$  and  $b$  is  $\delta_1$  "shifted" forward by one we see that in order to be able to prove a.s. convergence in Theorem 3.1 it is necessary that

$$(4.1) \quad \lim_{N \rightarrow \infty} X_{N+1} / \sum_1^N X_i = 0 \quad \text{a.s.}$$

Now by the generalized Borel-Cantelli Lemma (see [1, p. 96]) we know that (4.1) holds only if

$$(4.2) \quad \sum_1^\infty P \left[ X_{N+1} > \sum_1^N X_i | X_1, \dots, X_N \right] < \infty$$

almost surely. Now  $P[X_{N+1} > x] \sim x^{-\alpha}$  for  $x$  large by [3, p. 424], hence (4.2) is equivalent to the a.s. convergence of

$$(4.3) \quad \sum_{N=1}^\infty \left( \sum_{i=1}^N X_i \right)^{-\beta}$$

when  $\beta = \alpha$ . We will show that in fact (4.3) diverges when  $\beta = \alpha$ . (Note that by [9, p. 758] we know that (4.3) converges for  $\beta > \alpha$  and diverges for  $\beta < \alpha$ .)

To see that (4.3) diverges when  $\beta = \alpha$  we remember first that the series  $\sum_{N=2}^\infty (N \log(N))^{-1}$  diverges. We note also that

$$\begin{aligned} P \left[ \left( \sum_{i=1}^N X_i \right)^{-\alpha} > \frac{1}{N \log N} \right] &= P \left[ \sum_{i=1}^N X_i < (N \log N)^{1/\alpha} \right] \\ &= P [X_1 < (\log N)^{1/\alpha}] \end{aligned}$$

and that the last expression tends to 1 as  $N \rightarrow \infty$ . We conclude by Levy [7, Theorem 74.2, p. 297] that the series in (4.3) is a.s. divergent if  $\beta = \alpha$ .

#### REFERENCES

1. L. Breiman, *Probability*, Addison-Wesley, Reading, Mass., 1968. MR 37 #4841.
2. H. Cramér and M. R. Leadbetter, *Stationary and related stochastic processes. Sample function properties and their applications*, Wiley, New York, 1967. MR 36 #949.
3. W. Feller, *An introduction to probability theory and its applications*. Vol. II, 2nd ed., Wiley, New York, 1966. MR 35 #1048.
4. G. H. Hardy, J. E. Littlewood and G. Pólya, *Inequalities*, 2nd ed., Cambridge Univ. Press, New York, 1952. MR 13, 727.
5. M. Kanter and W. L. Steiger, *Regression and autoregression with infinite variance*, *Advances in Appl. Probability*, 6 (1974), 768–783.

6. J. Keilson, *Green's function methods in probability theory*, Griffin's Statistical Monographs and Courses, no. 17, Hafner, New York, 1965. MR 34 #2063.

7. P. Lévy, *Théorie de l'addition des variables aléatoires*, 2nd ed., Gauthier-Villars, Paris, 1954.

8. M. Loève, *Probability theory. Foundations. Random sequences*, 2nd rev. ed., University Ser. in Higher Math., Van Nostrand, Princeton, N. J., 1960. MR 23 #A670.

9. J. L. Miller, *A note on sums of independent random variables with infinite 1st moment*, Ann. Math. Statist. 38 (1967), 751–758.

DEPARTMENT OF MATHEMATICS, SIR GEORGE WILLIAMS UNIVERSITY, MONTREAL, QUEBEC, CANADA

*Current address:* School of Mathematics, University of New South Wales, Kensington, New South Wales, Australia 2033