Canad. Math. Bull. Vol. 16 (4), 1973.

ON THE EQUIVALENCE OF MODES OF CONVERGENCE[1]

BY
R. J. TOMKINS

1. **Introduction.** Let $(\Omega, \mathscr{F}, P)$ be a probability space. Let R denote the set of real numbers and $\mathscr{X}$ the set of all random variables defined on Ω. Throughout this work, random variables which differ only on a set of probability zero will be considered identical. EX represents, as usual, the expectation of X, $X \in \mathscr{X}$.

An event $A \in \mathscr{F}$ is called an *atom* if $P(A)>0$ and for any $B \subset A$ $(B \in \mathscr{F})$ either $P(B)=0$ or $P(B)=P(A)$. It is easy to show (see Loève's book [2] p. 100, for example) that any sample space can be written

$$\Omega = A \cup \bigcup_{k=1}^{\infty} A_k \tag{1}$$

where the events involved are disjoint, each A_k is either an atom or empty, and A has the property that, given any $B \in \mathscr{F}$ such that $B \subset A$ and any ε between 0 and $P(B)$, there exists $C \in \mathscr{F}$ such that $P(C)=\varepsilon$. It is easily demonstrated that random variables are constant on atoms.

Let $X, X_1, X_2, \ldots$ be in $\mathscr{X}$. The concepts of convergence of the sequence $\{X_n\}$, $n \geq 1$, to X almost certainly (denoted $X_n \xrightarrow{\text{a.c.}} X$), in probability ($X_n \xrightarrow{\text{p}} X$) and in the r^{th}-mean ($X_n \xrightarrow{\text{r}} X$) are well-known. Less familiar are the following two modes of convergence.

DEFINITION.

(i) Let $f:\mathscr{X} \to R$. Then the sequence $\{X_n\}$ is said to *converge to X in f* (denoted $X_n \xrightarrow{\text{f}} X$) if $f(X_n - X) \to 0$.

(ii) The sequence $\{X_n\}$ *converges completely* to X ($X_n \xrightarrow{\text{c}} X$) if

$$\sum_{n=1}^{\infty} P[|X_n - X| \geq \varepsilon] < \infty$$

for all $\varepsilon > 0$.

Type (i) is used by Thomasian [3]; his results will appear in section 2. Complete convergence was introduced by Hsu and Robbins [1]. By the Borel-Cantelli lemma, complete convergence implies convergence almost certainly. Lemma 1 provides a partial converse to this statement.

Received by the editors October 5, 1971, and, in revised form, August 23, 1972.
[1] The author is grateful to the National Research Council for financial support during the preparation of this work.

LEMMA 1. *Let* X, $X_n(n\geq 1)\in\mathscr{X}$. *If* $X_n\xrightarrow{\text{a.c.}}X$ *then there exists an increasing sequence* $\{n_k\}$ *of integers such that* $X_{n_k}\xrightarrow{\text{c}}X$.

Proof. For each m, $n=1, 2, \ldots$ define the events $A_{nm}=[|X_n-X|\geq 1/m]$. Since $X_n\xrightarrow{\text{a.c.}}X$ it follows that, for each $m\geq 1$,

$$\inf_{j\geq 1} P\left\{\bigcup_{n=j}^{\infty} A_{nm}\right\} = 0.$$

Let $n_1=1$. For $k>1$, let n_k be the least integer greater than n_{k-1} satisfying $P\{\bigcup_{n\geq n_k} A_{nk}\}\leq 2^{-k}$. Given $\varepsilon>0$, choose an integer $N>\varepsilon^{-1}$. Then

$$\sum_{k=N}^{\infty} P[|X_{n_k}-X|\geq\varepsilon]\leq\sum_{k=N}^{\infty} P(A_{n_k k})\leq\sum_{k=N}^{\infty} P\left\{\bigcup_{n\geq n_k} A_{nk}\right\}<2.$$

i.e. $X_{n_k}\xrightarrow{\text{c}}X$, as $k\to\infty$. Q.E.D.

The question to which this paper addresses itself is: what kind of restrictions have to be placed on Ω in order to insure that one type of convergence occurs if and only if another type occurs?

2. **The main result.** Clearly convergence in the r^{th} mean and convergence in f for some f are equivalent in any probability space; take $f(X)=E|X|^r$.

The next two theorems were proved by Thomasian [3].

THEOREM 1. *The following statements are equivalent:*

(i) *for some* $f:\mathscr{X}\to R$, $X_n\xrightarrow{\text{f}}X$ *if and only if (iff)* $X_n\xrightarrow{\text{a.c.}}X$ *for any* X_n, $X\in\mathscr{X}$

(ii) *for any* X_n, $X\in\mathscr{X}$, $X_n\xrightarrow{\text{a.c.}}X$ *iff* $X_n\xrightarrow{\text{p}}X$.

(iii) Ω *is the countable (possibly finite) union of disjoint atoms (i.e.* $A=\phi$ *in* (1)).

THEOREM 2. *The following are equivalent:*

(i) Ω *is the finite union of disjoint atoms.*

(ii) *for some* $f:\mathscr{X}\to R$, *convergence in* f *and convergence in probability are equivalent, and,*

$$|f(\alpha X)| = |\alpha|\cdot|f(X)| \quad \textit{for all } \alpha\in R,\ X\in\mathscr{X}.$$

The main result of the present work, theorem 3 below, extends theorem 2 to cover the remaining possible equivalences of convergence.

THEOREM 3. *The following are equivalent:*

(i) Ω *is the finite union of disjoint atoms.*

(ii) *for any* X, $X_n\in\mathscr{X}$, $X_n\xrightarrow{\text{p}}X$ *iff* $X_n\xrightarrow{\text{r}}X$ *for some (equivalently, all)* $r>0$.

(iii) *for any* X, $X_n\in\mathscr{X}$, $X_n\xrightarrow{\text{a.c.}}X$ *iff* $X_n\xrightarrow{\text{r}}X$ *for some (all)* $r>0$.

(iv) *for any* X, $X_n\in\mathscr{X}$, $X_n\xrightarrow{\text{c}}X$ *iff* $X_n\xrightarrow{\text{r}}X$ *for some (all)* $r>0$.

(v) *for some* $f:\mathscr{X}\to R$, *and any* X, $X_n\in\mathscr{X}$, $X_n\xrightarrow{\text{f}}X$ *iff* $X_n\xrightarrow{\text{c}}X$.

(vi) *for any* X, $X_n\in\mathscr{X}$, $X_n\xrightarrow{\text{p}}X$ *iff* $X_n\xrightarrow{\text{c}}X$.

(vii) *for any* X, $X_n\in\mathscr{X}$, $X_n\xrightarrow{\text{a.c.}}X$ *iff* $X_n\xrightarrow{\text{c}}X$.

https://doi.org/10.4153/CMB-1973-093-7 Published online by Cambridge University Press

Moreover, the function f in (v) *has the property that* $f(X)=0$ *iff* $X=0$ *a.c. and f may, without loss of generality, be assumed to be non-negative.*

3. **Proof of Theorem 3.** Since $X_n \to X$ iff $X_n - X \to 0$ for any of the five convergence modes under consideration, there is no harm in assuming throughout the proof that $X \equiv 0$.

To prove theorem 3, it will be shown that each statement in the theorem implies the following one and that (vii) implies (i).

(i) implies (ii): Suppose $\Omega = \bigcup_{k=1}^N A_k$, where the A_k's are disjoint atoms. Suppose $X_n \xrightarrow{\text{p}} 0$. Since random variables are constant on atoms, we can find constants c_{nk}, for each $n \geq 1$ and $k \leq N$, such that $X_n = c_{nk}$ on A_k. Define $c_n = \max_{k \leq N} |c_{nk}|$. By theorem 1, $X_n \xrightarrow{\text{a.c.}} 0$. Hence, given $\varepsilon > 0$, there exists an integer $N_k (k \leq N)$ such that $|c_{nk}| < \varepsilon$ if $n \geq N_k$. If $M = \max_{k \leq N} N_k$, $c_n < \varepsilon$ for $n \geq M$; i.e. $c_n \to 0$ as $n \to \infty$.

Now, for any $r > 0$, $E|X_n|^r \leq c_n^r \to 0$. So $X_n \xrightarrow{r} 0$ for any $r > 0$; i.e. (ii) holds.

(ii) implies (iii). Suppose (ii) holds; let $r > 0$. Letting $f(X) = E|X|$, it is clear from (ii) that statement (ii) of theorem 2 holds. This, in turn, implies that (i) holds. But then theorem 1 applies, yielding the fact that $X_n \xrightarrow{\text{p}} 0$ iff $X_n \xrightarrow{\text{a.c.}} 0$. This equivalence, together with (ii), gives (iii).

(iii) implies (iv). If (iii) holds and $X_n \xrightarrow{\text{c}} 0$, then trivially $X_n \xrightarrow{\text{r}} 0$ for all $r > 0$. If (iii) holds for some $r > 0$, let $f(X) = (E|X|^r)^{1/r}$. Then $X_n \xrightarrow{\text{f}} 0$ iff $X_n \xrightarrow{\text{a.c.}} 0$ so that it follows by theorem 1 that $X_n \xrightarrow{\text{a.c.}} 0$ iff $X_n \xrightarrow{\text{p}} 0$. Hence $X_n \xrightarrow{\text{f}} 0$ iff $X_n \xrightarrow{\text{p}} 0$ and f satisfies (ii) of theorem 2, implying that (i) holds, say $\Omega = \bigcup_{k=1}^N A_k$. Now suppose $X_n \xrightarrow{\text{r}} 0$ for some $r > 0$. By (iii), $X_n \xrightarrow{\text{a.c.}} 0$. Using an argument given earlier in this proof, one can prove that given $\varepsilon > 0$ there exists $M = M_\varepsilon$ such that $|X_n| < \varepsilon$ for all $n \geq M$. Then

$$\sum_{n=1}^{\infty} P[|X_n| \geq \varepsilon] = \sum_{n=1}^{M} P[|X_n| \geq \varepsilon] < \infty.$$

So $X_n \xrightarrow{\text{c}} 0$.

(iv) implies (v): clear if one takes $f(X) = E|X|^r$.

(v) implies (vi): Suppose $X_n \xrightarrow{\text{p}} 0$ but $X_n \overset{\text{c}}{\not\to} 0$. By (v), $f(X_n) \not\to 0$, i.e. for some $\varepsilon > 0$ and subsequence of integers $n_n \uparrow \infty$,

$$|f(X_{n_k})| \geq \varepsilon, \quad k \geq 1. \tag{2}$$

Since $X_{n_k} \xrightarrow{\text{p}} 0$, there is a subsequence of $\{X_{n_k}\}$ converging to zero almost certainly. By lemma 1 that subsequence has a subsequence converging completely to zero, say $X_{n_{k_j}} \xrightarrow{\text{c}} 0$ as $j \to \infty$. Hence, by (iv), $f(X_{n_{k_j}}) \to 0$ as $j \to \infty$ which contradicts (2).

https://doi.org/10.4153/CMB-1973-093-7 Published online by Cambridge University Press

That (vi) implies (vii) is obvious. (vii) implies (i): Suppose (vii) holds and Ω is given by (1). If $A \neq \phi$, then choose a decreasing sequence of events $A \supset B_1 \supset B_2 \supset \cdots$ such that $P(B_n) = \min(1/n, P(A))$.) For each n, define the random variable,

$$X_n = \begin{cases} 1 & \text{on } B_n \\ 0 & \text{elsewhere.} \end{cases}$$

It is easily shown that $X_n \xrightarrow{\text{a.c.}} 0$. But, if $1/M < PA$, then

$$\sum_{n=M}^{\infty} P[|X_n| \geq 1] = \sum_{n=M}^{\infty} n^{-1} = \infty,$$

so $X_n \xrightarrow{\text{c}} 0$ is false, contradicting (vii). Thus $A = \phi$. Suppose atoms $A_1, A_2, \ldots$ exist such that $\Omega = \bigcup_{k=1}^{\infty} A_k$. For each $k \geq 1$ let N_k=first integer greater than $1/P(A_k)$. For each $n \geq 1$ define the random variable Y_n as follows: if $\omega \in A_k$ put $Y_n(\omega) = 2^{N_k - n}$. (Note that, given any element ω of Ω, a unique index k ($k \geq 1$) exists such that $\omega \in A_k$.) Clearly $Y_n \xrightarrow{\text{a.c.}} 0$. So, by (vii), $Y_n \xrightarrow{\text{c}} 0$. But

$$\begin{aligned}
\sum_{n=1}^{\infty} P[|Y_n| \geq 1] &= \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} P\{[Y_n \geq 1] \cap A_k\} \\
&= \sum_{k=1}^{\infty} \sum_{n=1}^{\infty} P\{[2^{N_k} \geq 2^n] \cap A_k\} \\
&= \sum_{k=1}^{\infty} \sum_{n=1}^{N_k} P(A_k) = \sum_{k=1}^{\infty} N_k P(A_k) > \sum_{k=1}^{\infty} 1 = \infty,
\end{aligned}$$

contradicting $Y_n \xrightarrow{\text{c}} 0$. Hence Ω can only be a finite union of atoms.

It remains to prove the last statement of the theorem. Suppose (v) holds for some f; then (v) also is valid for $|f|$.

Define $X_n = 0$ for all n. $X_n \xrightarrow{\text{c}} 0$ so $f(X_n) = f(0) = 0$. If $f(Y) = 0$ for some $Y \in \mathscr{X}$ where $P[Y \neq 0] > 0$, then $f(Y_n) \to 0$ if $Y_n = Y$ for each n. But then $Y_n \xrightarrow{\text{c}} 0$ which is clearly false.

REMARK. If (v) of theorem 3 holds for some f, then obviously (ii) of theorem 2 holds—but not necessarily for the same function f.

For example, if $\mathscr{F} = \{\phi, \Omega\}$ then all random variables are constants. Let $f(X) = X^2$. (v) of theorem 3 holds for this f, but f does not satisfy (ii) of theorem 2.

4. **On norms and metrics.** It is well-known that convergence in the metric $d(X, Y) \equiv \dfrac{E|X-Y|}{1+E|X-Y|}$ is equivalent to convergence in probability in any probability space.

Thomasian [3] proved that a metric exists for $\mathscr{X}$ such that convergence in the metric and convergence almost certainly are equivalent if and only if Ω is a countable union of disjoint atoms. His result is extended by the following theorem.

THEOREM 4. *The following are equivalent:*

(i) Ω *is the finite union of disjoint atoms.*

(ii) *a metric d exists for* $\mathscr{X}$ *such that* $X_n \xrightarrow{c} X$ *iff* $d(X_n, X) \to 0$ *as* $n \to \infty$.

Proof. (i) implies (ii): If (i) holds then, by theorem 3(iii), $X_n \xrightarrow{c} X$ iff $E|X_n - X| \to 0$. (ii) holds with $d(X, Y) \equiv E|X-Y|$.

(ii) implies (i): Suppose d is a metric for $\mathscr{X}$; define $f(X)=d(X, 0)$. Then $X_n \xrightarrow{c} X$ iff $X_n - X \xrightarrow{c} 0$ iff $d(X_n - X, 0) \to 0$ iff $X_n \xrightarrow{f} X$. Thus theorem 3(iv) holds, which is equivalent to (i).

Thomasian further exhibited that a norm on $\mathscr{X}$ exists such that convergence in the norm and convergence in probability always occur together iff Ω is the finite union of disjoint atoms. Theorem 5 shows that Thomasian's result remains valid when the phrase "in probability" is replaced by either "almost certainly" or "completely".

THEOREM 5. *The following are equivalent:*

(i) Ω *is a finite union of disjoint atoms.*

(ii) *a norm exists for* $\mathscr{X}$ *such that convergence in the norm is equivalent to convergence completely.*

(iii) *a norm exists for* $\mathscr{X}$ *such that convergence in the norm is equivalent to convergence almost certainly.*

Proof. If (i) holds, let $f(X)=E|X|$. This is clearly a norm for $\mathscr{X}$. (ii) and (iii) follow by theorem 3. If (iii) holds, then (i) of theorem 1 holds. Hence, using theorem 1(ii), convergence in the norm and convergence in probability are equivalent.

But, as previously remarked, Thomasian proved that equivalence of convergence in the norm and convergence in probability occurs iff (i) holds. If (ii) holds then (iv) of theorem 3 is satisfied, so (i) holds.

REFERENCES

1. P. L. Hsu and Herbert Robbins, *Complete convergence and the law of large numbers*, Proc. Acad. Sci. U.S.A. **33** (1947), 25–31.

2. M. Loève, *Probability theory*, 3*rd edition*, Van Nostrand, Princeton, N.J., 1963.

3. A. J. Thomasian, *Metrics and norms on spaces of random variables*, Ann. Math. Statist. **28** (1957), 512–514.

UNIVERSITY OF SASKATCHEWAN, REGINA CAMPUS
REGINA, SASKATCHEWAN

https://doi.org/10.4153/CMB-1973-093-7 Published online by Cambridge University Press