The Annals of Probability
1989, Vol. 17, No. 1, 353–356

A CONVERGENCE PROPERTY FOR CONDITIONAL EXPECTATION[1]

By Aurel Cornea and Peter A. Loeb

Katholische Universität Eichstätt and University of Illinois

Convergence properties are obtained for repeated applications of the operator $f \to |f - E(f)|$, where E denotes conditional expectation. If, for example, E is the integral with respect to a probability measure P, $f \in L^\infty(P)$ and $T(f) = |f - E(f)|$, then $T^n(f)$ converges to 0 in $L^\infty(P)$ and $\Sigma T^n(f)$ converges in $L^1(P)$.

Fix a probability space $(X, \mathscr{B}, \mu)$ and let $\mathscr{A}$ be a sub-σ-algebra of $\mathscr{B}$. Let $L(\mathscr{A}, \mathscr{B})$ denote the set of functions f on X such that $f \in L^1(\mathscr{B})$ and there is a function H_f in $L^1(\mathscr{A})$ with $|f| \le H_f$ almost everywhere. Given $f \in L(\mathscr{A}, \mathscr{B})$, let $E(f)$ denote the conditional expectation with respect to $\mathscr{A}$ and let $T(f)$ denote the function $|f - E(f)|$. Further, let $M(f)$ be the greatest lower bound of the "functions" in $L^1(\mathscr{A})$ which dominate f almost everywhere. As usual, we are really working with equivalence classes of functions and with relations that hold almost everywhere with respect to $(X, \mathscr{B}, \mu)$. For example, $M(f)$ dominates f almost everywhere and has the smallest integral of the family of dominating functions in $L^1(\mathscr{A})$. If $\mathscr{A}$ is the trivial σ-algebra $\{X, \phi\}$, then $M(f)$ is the essential supremum of f.

Theorem. *Given $f \in L(\mathscr{A}, \mathscr{B})$, the sequence $M(T^n(f))$ converges to 0 almost everywhere and the series $\Sigma_1^\infty T^n(f)$ converges in $L^1(\mathscr{B})$; in fact, $\Sigma_1^\infty E(T^n(f)) \le 4M(|f|)$. If $\mathscr{A}$ is the trivial σ-algebra, then $T^n(f)$ converges to 0 in $L^\infty(\mathscr{B})$.*

Remark. Let $\{\mathscr{A}_n \colon n \in \mathbb{N}\}$ be an arbitrary sequence of sub-σ-algebras of $\mathscr{B}$. For each $n \in \mathbb{N}$, let E_n denote conditional expectation with respect to $\mathscr{A}_n$, T_n be the mapping $f \to |f - E_n(f)|$ and V_n be the mapping $T_n \circ T_{n-1} \circ \cdots \circ T_1$. Testing the preceding results with L^2 methods, J. L. Doob in private correspondence has shown that $V_n(f)$ converges to 0 almost everywhere when $f \in L^2(\mathscr{B})$. On the other hand, if $\{\mathscr{A}_n \colon n \in \mathbb{N}\}$ is an increasing sequence of sub-σ-algebras in $\mathscr{B}$, one can show with a slight modification of our proof that $\Sigma_1^\infty E_n(V_n(f)) \le 4M_1(|f|)$, where M_1 denotes the greatest lower bound in $L^1(\mathscr{A}_1)$ and $f \in L(\mathscr{A}_1, \mathscr{B})$.

Proof of the theorem. To prove the theorem we need some preliminary results. First we note that T is subadditive, i.e., for all $f, g \in L(\mathscr{A}, \mathscr{B})$, $T(f + g) \le T(f) + T(g)$. For each $f \in L(\mathscr{A}, \mathscr{B})$, $T(f)$ is again in $L(\mathscr{A}, \mathscr{B})$; we set $m(f) = -M(-f)$. Given $h \in L^1(\mathscr{A})$, $T(f + h) = T(f)$. If, moreover, h

Received September 1987.
[1]Work supported in part by NSF Grant DMS-84-01294.
AMS 1980 *subject classifications*. Primary 28A20, 60F25.
Key words and phrases. Conditional expectation, L^1 convergence.

takes only the values 1 and -1, then $T(hf) = T(f)$. If $H = M(|f|)$, then $f + H$ is nonnegative almost everywhere and $T(f + H) = T(f)$; also note that $2M(f + H) \leq 4H = 4M(|f|)$. Thus for the proof we need only consider a nonnegative function $f \in L(\mathscr{A}, \mathscr{B})$, but we replace $4M(|f|)$ with $2M(f)$ in the statement of the theorem. For such an f we have $M(T(f)) \leq M(f)$.

We define two new operators s and S on $L(\mathscr{A}, \mathscr{B})$. The first is obtained by setting $s(g)$ equal to the characteristic function of the set $U(g)$ in $\mathscr{A}$ where $(M(g) + m(g))/2 \geq E(g)$. (Where necessary, we change the value $+\infty$ or $-\infty$ to 0.) The second operator is $S = E + sT - (1 - s)T$. Fix $g \in L(\mathscr{A}, \mathscr{B})$. Since

$$(1 - s(g))m(g) + s(g)E(g) \leq S(g) \leq s(g)M(g) + (1 - s(g))E(g)$$

and $m(g) \leq E(g) \leq M(g)$, we have

$$m(g) \leq m(S(g)) \leq M(S(g)) \leq M(g).$$

On the subset of $U(g)$ where $g \geq E(g)$, $S(g) = g$. On the subset of $U(g)$ where $g < E(g)$, we have

$$0 \leq S(g) - g = 2(E(g) - g) \leq 2(m(S(g)) - g) \leq 2(m(S(g)) - m(g)).$$

Thus on the entire set $U(g)$, $|S(g) - g| \leq 2(m(S(g)) - m(g))$. On the complement of $U(g)$ we have by similar calculations the inequality $|S(g) - g| \leq 2(M(g) - M(S(g)))$, so on all of X,

$$|S(g) - g| \leq 2(M(g) - M(S(g)) + m(S(g)) - m(g)).$$

It follows that

$$M(|S(g) - g|) \leq 2(M(g) - M(S(g)) + m(S(g)) - m(g)).$$

To simplify our notation, we fix a nonnegative $f \in L(\mathscr{A}, \mathscr{B})$ and set $f_0 = f$ and $f_n = S^n(f)$ for $n \geq 1$. We further set $h_n = 2s(f_n) - 1$ for all $n \geq 0$. It follows from the preceding calculations that

$$\sum_0^\infty |f_{n+1} - f_n| \leq \sum_0^\infty M(|f_{n+1} - f_n|) \leq 2(M(f) - m(f)) \leq 2M(f).$$

In particular, f_n converges almost everywhere. From the definition we have $f_{n+1} = E(f_n) + h_n T(f_n)$. Using this equality and the properties of T stated at the beginning of the proof, one can show inductively that

$$T^{n+1}(f) = T(f_n) = h_n(f_{n+1} - E(f_n)) = |f_{n+1} - E(f_n)|.$$

Thus

$$E(T^{n+1}(f)) = h_n E(f_{n+1} - f_n) = |E(f_{n+1} - f_n)| \leq E(|f_{n+1} - f_n|) \\ \leq M(|f_{n+1} - f_n|),$$

so,

$$\sum_1^\infty E(T^n(f)) \leq \sum_0^\infty M(|f_{n+1} - f_n|) \leq 2M(f).$$

Now we note that

$$\int\sum_1^\infty T^n(f) = \sum_1^\infty \int E(T^n(f)) \leq 2\int M(f).$$

Hence $\sum_1^\infty T^n(f)$ is finite almost everywhere, so $T^n(f)$ converges to 0 almost everywhere. Let F denote the limit of the f_n's. Since $f_{n+1} = E(f_n) + h_n T^{n+1}(f)$, $F = E(F)$ a.e. by the dominated convergence theorem for conditional expectations. We may therefore assume that F is $\mathscr{A}$-measurable. By the subadditivity of M,

$$M\big(T^{n+1}(f)\big) = M\big(|f_{n+1} - E(f_n)|\big) \leq M\big(|f_{n+1} - F|\big) + |F - E(f_n)|.$$

It is also easy to see that

$$M(|f_n - F|) \leq \sum_n^\infty M(|f_{k+1} - f_k|).$$

Hence $\lim_{n\to\infty} M(|f_n - F|) = 0$, so $\lim_{n\to\infty} M(T^n(f)) = 0$ almost everywhere. □

EXAMPLE 1. Let $X = \{a, b\}$. Let $\mathscr{B}$ be the power set of X and let $\mathscr{A} = \{X, \phi\}$. Also let $\mu(\{a\}) = \alpha$ and $\mu(\{b\}) = \beta$ with $0 < \alpha < 1$. Let D and V denote the operators on $L^\infty(\mathscr{B})$ defined by setting $D(f) = f - E(f)$ and $V(f) = |f|$. Note that $T = V \circ D$; we set $W = D \circ V$. It follows that $T \circ V = V \circ W$ and for any $n \geq 1$ and any f, $T^{n+1}(f) = V(W^n(D(f)))$.

Fix f with $f(a) = 1$ and $f(b) = 0$. Now $D(f)$ takes the values $u_0 = \beta$ at a and $v_0 = -\alpha$ at b, whence $\alpha|u_0| = \beta|v_0|$. For $n \geq 1$, let $W^n(D(f))$ take the values u_n at a and v_n at b. If $n \geq 0$ and $\alpha|u_n| = \beta|v_n|$, then

$$u_{n+1} = |u_n| - \big(\alpha|u_n| + \beta|v_n|\big) = \beta\big(|u_n| - |v_n|\big) = |u_n|(1 - 2\alpha),$$
$$v_{n+1} = |v_n| - \big(\alpha|u_n| + \beta|v_n|\big) = \alpha\big(|v_n| - |u_n|\big) = |v_n|(1 - 2\beta)$$

and

$$\alpha|u_{n+1}| = \beta|v_{n+1}|.$$

By iteration we have for all $n \geq 0$, $T^{n+1}(f)(a) = |u_n| = \beta|1 - 2\alpha|^n$ and $T^{n+1}(f)(b) = |v_n| = \alpha|1 - 2\beta|^n$, so

$$\sum_1^\infty E(T^n(f)) = \alpha\beta\left(\frac{1}{1 - |1 - 2\alpha|} + \frac{1}{1 - |1 - 2\beta|}\right) = \max(\alpha, \beta).$$

EXAMPLE 2. Let $(X_i, \mathscr{B}_i, \mu_i)$, $i \in \mathbb{N}$, be a sequence of probability spaces each of the type of Example 1 with $X_i = \{a_i, b_i\}$, $\mu_i(\{a_i\}) = \alpha_i$ and $\mu_i(\{b_i\}) = \beta_i$. Let $(X, \mathscr{B})$ be the direct sum of the spaces $(X_i, \mathscr{B}_i)$ and let $\mu = \sum_1^\infty 2^{-i}\mu_i$. Let $\mathscr{A}$ be the σ-algebra generated by the sets X_i. Each of the operations E, T and M are computed independently on each subset X_i of X. Take f with $f(a_i) = 1$ for each i and $f(b_i) = 0$ for each i. Then $T^{n+1}(f)(a_i) = \beta_i|1 - 2\alpha_i|^n$. If for all i we choose $\alpha_i < \frac{1}{2}$ so that $(1 - 2\alpha_i)^i = \frac{1}{2}$, we have $T^{n+1}(f)(a_n) \geq \frac{1}{4}$. In this case, the sequence $T^n(f)$ does not converge in $L^\infty(\mathscr{B})$.

Acknowledgment. The authors are indebted to Professor J. L. Doob for helpful conversations.

KATHOLISCHE UNIVERSITÄT EICHSTÄTT
OSTENSTRASSE 26-28
D-8078 EICHSTÄTT
FEDERAL REPUBLIC OF GERMANY

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF ILLINOIS
1409 WEST GREEN STREET
URBANA, ILLINOIS 61801