A Uniform Operator Ergodic Theorem

by

E. M. Klimko

Department of Statistics
Division of Mathematical Sciences
Mimeograph Series No. 154
April 1968

A Uniform Operator Ergodic Theorem

by

E.M. Klimko

1. Introduction. The purpose of this note is to prove a uniform operator ergodic theorem for mean convergence of differences of right continuous stochastic processes. Our result contains a difference version of the Glivenko-Cantelli theorem for infinite invariant measures. We also state a pointwise convergence theorem valid in the presence of a positive fixed point, which generalizes a result of Burke [1].

Let $(\Omega, \textcircled{A}, P)$ be a probability space; let L_1 be the class of integrable functions on $(\Omega, \textcircled{A}, P)$ and let L_1^+ be the class of non-negative integrable functions and let T be a Markovian operator mapping L_1 into itself. A set A is closed if for each $f \epsilon L_1$, $f = 0$ on A^c implies that $Tf = 0$ on A^c. The class of closed sets forms the invariant sigma field $\textcircled{I}$. T is ergodic if $\textcircled{I}$ is trivial. Let T_∞ denote the (formal) operator $I + T + T^2 + \ldots$. Hopf's decomposition states that $\Omega = C + D$ where for every $f \epsilon L_1^+$, $T_\infty f = 0$ or ∞ on C and $T_\infty f < \infty$ on D. If $\Omega = C$, T is conservative. We now state an ergodic theorem required in the sequel. We use the notation:

$$E(f) = \int_\Omega f \, d\mu. \tag{1.1}$$

Theorem 1.1. Let T be a conservative ergodic Markovian operator. If $f \epsilon L_1$ and $E(f) = 0$, then

$$(1.2) \qquad n^{-1} (f + Tf + \dots + T^{n-1} f)$$

converges to zero in the L_1 topology. Theorem 1.1 was obtained by Sucheston [6] and independently by Krengel [5].

2. Main Results. Let $X(w,t)$, $Y(w,t)$ be left continuous nondecreasing stochastic processes on $\Omega \times R$ such that for each $t \epsilon R$ $EX(w,t) < \infty$ and $E\,Y(w,t) < \infty$. We will omit w in $X(w,t)$ for simplicity. Set $X_n(t) = T^n X(t)$, $Y_n(t) = T^n Y(t)$ for $n = 0,1,\dots$. We may and do assume that $X_n(t)$ and $Y_n(t)$ are chosen in such a way that outside a null set N independent of t and n, they are nondecreasing and left continuous functions of t. Such a choice is possible by a regularization procedure as used in constructing regular conditional probabilities.

In [1], it was shown that if T is generated by a point transformation which preserves a finite measure, then the cesaro averages of $T^n X(t)$ converge almost everywhere uniformly with respect to t on compact intervals. Here we show that when suitably normalized, $X(t)$ and $Y(t)$ behave similarly in the mean, uniformly on a rectangle.

Theorem 2.1. Let $0 < c_1 \leq c_2 < \infty$, $0 < d_1 \leq d_2 < \infty$; let

$$(2.1) \qquad C = \{t: c_2 \geq E\,X(t) \geq c_1\} \quad D = \{t: d_2 \geq E\,Y(t) \geq d_1\}.$$

Let $B = C \times D$ and

$$(2.2) \qquad \Delta_n = \sup_{(s,t)\epsilon B} \left| \frac{1}{n} \sum_{k=0}^{n-1} \frac{X_k(s)}{EX(s)} - \frac{Y_k(t)}{EY(t)} \right|.$$

Then Δ_n converges to zero in the L_1 topology.

Proof. We may and do assume that $c_1 = \inf_{t\in C} E\,X(t)$, $c_2 = \sup_{t\in C} E\,X(t)$, $d_1 = \inf_{t\in D} E\,Y(t)$, $d_2 = \sup_{t\in D} E\,Y(t)$. For each fixed integer m and each $j = 1,2,\ldots m-1$, we let s_{mj}, t_{mj} be the smallest real numbers such that:

$$(2.3)\qquad \begin{cases} E\,X(s_{mj}) \leq c_1 + j(c_2 - c_1)\,/\,m \leq E\,X(s_{mj}+0) \\ E\,Y(t_{mj}) \leq d_1 + j(d_2 - d_1)\,/\,m \leq E\,Y(t_{mj}+0) \end{cases}$$

Further set $s_{mo} = \inf C$, $s_{mm} = \sup C$, $t_{m0} = \inf D$, $t_{mm} = \sup D$. For each pair $(s,t)\in B$, we define

$$(2.4)\qquad \delta_n(s,t) = X_n(s)/EX(s) - Y_n(t)/EY(t)$$

It follows from theorem 1.1 applied to $\delta_0(s,t)$ that for fixed s,t, $\delta_n(s,t)$ converges cesaro in the L_1 topology to zero. Since positive linear operators are order preserving, for $s_{m,i-1} < s \leq s_{mi}$, $t_{m,j-1} < t \leq t_{mj}$ we have

$$(2.5)\qquad \begin{cases} \dfrac{X_k(s_{m,i-1}+0)}{EX(s_{mi})} \leq \dfrac{X_k(s)}{EX(s)} \leq \dfrac{X_k(s_{m,i})}{EX(s_{m,i-1}+0)} \\[2ex] \dfrac{Y_k(t_{m,j-1}+0)}{EY(t_{m,j})} \leq \dfrac{Y_k(t)}{EY(t)} \leq \dfrac{Y_k(t_{m,j})}{EY(t_{m,j-1}+0)} \end{cases}$$

and

$$(2.6) \qquad \frac{X_k(s_{m,i-1}+0)}{EX(s_{mi})} - \frac{Y_k(t_{m,j})}{EY(t_{m,j-1}+0)} \leq \delta_k(s,t) \leq$$

$$\frac{X_k(s_{mi})}{EX(s_{m,i-1}+0)} - \frac{Y_k(t_{m,j-1}+0)}{EY(t_{m,j})}$$

From (2.3) it follows that

$$EX(s_{mi})/ EX(s_{m,i-1}+0) \leq 1 + c/m$$

$$EY(t_{mj-1}+0) / EY(t_{mj}) \geq 1 - d/m$$

where $c = (c_2 - c_1) / c_1$, $d = (d_2 - d_1) / d_1$.

Therefore

$$(2.7) \qquad \delta_k(s,t) \leq (1 + c/m)\, \delta_k(s_{mi}, t_{m,j-1}+0)$$

$$+(c/m + d/m)\, Y_k(t_{m,j-1}+0)/EY(t_{m,j-1}+0).$$

By similar arguments, we obtain a lower bound

$$(2.8) \qquad -\delta_k(s,t) \leq -(1-c/m)\, \delta_k(s_{m,i-1}+0, t_{mj})$$

$$+ (c/m + d/m)\, Y_k(t_{m,j-1}+0)/EY(t_{m,j-1}+0).$$

Since T is Markovian, the integral of $Y_k(s_{m,j-1}+0)/EY(s_{m,j-1}+0)$ is one and it follows that

$$(2.9) \qquad E|\Delta_n| \leq E \max (\Delta_n^{(1)}, \Delta_n^{(2)}) + (c + d)/m$$

where

$$(2.10) \begin{cases} \Delta_n^{(1)} = (1 + c/m) \max\limits_{\substack{0 \leq i \leq m \\ 0 \leq j \leq m}} \frac{1}{n} \sum\limits_{k=0}^{n-1} \delta_k (s_{mi}, t_{m,j-1} + 0) \\ \Delta_n^{(2)} = -(1-c/m) \max\limits_{\substack{0 \leq i \leq m \\ 0 \leq j \leq m}} \frac{1}{n} \sum\limits_{k=0}^{n-1} \delta_k (s_{m,i-1} + 0, t_{mj}) \end{cases}$$

Each of the terms over which max is taken in $\Delta_n^{(1)}$ and $\Delta_n^{(2)}$ converge to zero in the L_1 topology by theorem 1.1.

Therefore,

$$(2.11) \qquad \lim_n \sup \int |\Delta_n| \leq (c + d) / m$$

and since m is arbitrary, convergence in the L_1 topology follows.

If the operator T admits of a fixed point $f \in L_1^+$, we may obtain pointwise convergence. In this case, the role of theorem 1.1. may be played by Hopf's operator ergodic theorem [2].

Theorem 2.2. Let $T1 = 1$; $0 \leq c_1 < c_2 < \infty$; $C = \{t: c_1 < E\, X(t) < c_2\}$; and

$$(2.12) \qquad \Delta_n = \sup_{t \in C} \frac{1}{n} \left| \sum_{k=0}^{n-1} X_k(t) - E\, X(t) \right|.$$

Then for almost every $w \in \Omega$,

(2.13) $$\lim_{n \to \infty} \Delta_n = 0.$$

The proof of this theorem is similar to theorem 2.2 and is omitted. For the next theorem, we permit P to be sigma finite on Ⓐ. Let τ be a measure preserving, conservative, ergodic point transformation. τ generates a Markovian operator T by means of the relation $Tf = f \circ \tau$. This correspondence preserves the notions of ergodicity and conservativity of an operator. Let X_o, Y_o be fixed real-valued measurable functions on Ω and for $n=1,2,\ldots$, let $X_n = X_o \circ \tau^n$, $Y_n = Y_o \circ \tau^n$. If s,x,t,y are extended real numbers, let

(2.14) $$F_n^s(x) = 1_{(s,x)} \circ X_n, \quad G_n^t(y) = 1_{(t,y)} \circ Y_n \quad n=0,1\ldots$$

and

(2.15) $$F^s(x) = E(F_o^s(x)), \quad G^t(y) = E(G_o^t(y)).$$

Theorem 2.1 contains the following difference version of the Glivenko-Cantelli theorem for infinite invariant measures. A ratio version of this theorem was proved in [3].

<u>Theorem 2.3</u>. Let $s,t \in \overline{R}$ (extended real line). Let C and D be sets in $\overline{R}$ such that for some positive constants c_1, c_2, d_1, d_2

(2.16) $$C = \{x: c_2 \geq F^s(x) \geq c_1\}$$

$$D = \{y: d_2 \geq G^t(y) \geq d_1\}.$$

Let $B = C \times D$ and

$$(2.17) \qquad \Delta_n = \sup_{(x,y)\epsilon B} \frac{1}{n} \left| \sum_{i=0}^{n-1} \frac{F_i^s(x)}{F^s(x)} - \frac{G_i^t(y)}{G^t(y)} \right| .$$

Then Δ_n converges to 0 in the L_1 topology.

3. The non ergodic case. With suitable modifications, Theorems 2.1 and 2.2 remain valid even though the invariant σ-field Ⓘ is non-trivial. Theorem 1.1 was actually proved under the weaker condition $E(f \mid Ⓘ) = 0$. Therefore, we may now state Theorem 2.1 valid in the case when Ⓘ is not trivial. This theorem is based on an idea of Tucker [7].

Theorem 3.1. Let $c_1(w)$, $c_2(w)$, $d_1(w)$, $d_2(w)$ be Ⓘ-measurable; let

$$(3.1) \qquad \begin{aligned} C &= \{t: c_2 \geq E(X_o(t) \mid Ⓘ) \geq c_1) \\ D &= \{t: d_2 \geq E(Y_o(t) \mid Ⓘ) \geq d_1), \end{aligned}$$

the inequalities holding except on a null set N independent of t. Let $B = C \times D$ and

$$(3.2) \qquad \Delta_n = \sup_{(s,t)\epsilon B} \frac{1}{n} \left| \sum_{k=0}^{n-1} \frac{X_k(s)}{E(X_o(s) \mid Ⓘ)} - \frac{Y_k(t)}{E(Y_o(t) \mid Ⓘ)} \right|$$

Then $\lim_{n\to\infty} \int |\Delta_n| = 0$.

Proof. We merely sketch the proof since it is similar to theorem 2.1.

We may and do assume that

$$c_2 = \sup_{t\in C} E\,(X(t)\,|\,\textcircled{I}) \qquad c_1 = \inf_{t\in C} E\,(X(t)\,|\,\textcircled{I})$$

$$d_2 = \sup_{t\in D} E\,(Y(t)\,|\,\textcircled{I}) \qquad d_1 = \inf_{t\in D} E\,(Y(t)\,|\,\textcircled{I})$$

For each fixed integer m and each $j = 0,1,\ldots m-1$, we let $s_{mj}(w)$, $t_{mj}(w)$ be the smallest real numbers such that

$$(3.4)\quad \begin{cases} E\,(X(s_{mj})\,|\,\textcircled{I}) \leq c_1 + j(c_2 - c_1)/m \leq E\,(X(s_{mj}+0)\,|\,\textcircled{I}) \\ E\,(Y(t_{mj})\,|\,\textcircled{I}) \leq d_1 + j(d_2 - d_1)/m \leq E\,(Y(t_{mj}+0)\,|\,\textcircled{I}). \end{cases}$$

The functions s_{mj}, t_{mj} are measurable on the sigma field generated by $\textcircled{I}$ and are ordered: $s_{mj-1} \leq s_{mj}$, $t_{mj-1} \leq t_{mj}$, $j = 0,\ldots,m$. The arguments in the proof of theorem 2.1 apply with the preceeding changes.

Theorem 2.3 also extends to the non ergodic case except that the conditional expectation is not defined if $\textcircled{I}$ contains an atom of infinite measure. We may however, compute conditional expectation with respect to an equivalent probability measure (see [4]).

4. <u>Acknowledgement</u>. The author would like to thank L. Sucheston for comments and suggestions.

References

(1) Burke, G. (1965) A uniform ergodic theorem. Ann. Math. Statist. 36, 1853-1858.

(2) Hopf, E. (1954) The general temporally discrete Markov Process. J. Rat. Mech. and Anal. 3, 13-45.

(3) Klimko, E.M. (1967) On the Glivenko-Cantelli theorem for infinite invariant measures. Ann. Math. Statist. 38, 1273-1277

(4) Klimko, E.M. and Sucheston, L. An operator ergodic theorem for sequences of functions, to appear.

(5) Krengel, U. (1966) On the Global limit behavior of Markov chains and general non singular Markov processes. Z. Wahrscheinlichkeitstheorie. Verw. Geb. 6, 302-316.

(6) Sucheston, L. (1967) On the ergodic theorem for positive operators I. Z. Wahrscheinlichkeitstheorie. Verw. Geb. 8, 1-11.

(7) Tucker, H. G. (1959) A generalization of the Glivenko-Cantelli theorem. Ann. Math. Statist. 30, 828-830.