A UNIFORM ERGODIC THEOREM[1]

BY GEORGE BURKE

State University of Iowa

1. Introduction. Let $\{X_n\}$ be a sequence of independent random variables with common distribution function $F(t)$, and let $F_n(t, \omega)$ be the nth empirical distribution function of the sequence. Then the Glivenko-Cantelli theorem ([3], p. 20), states that for almost all ω, $F_n(t, \omega)$ converges to $F(t)$ uniformly in t. In [4], Tucker has shown that even if $\{X_n\}$ is only strictly stationary $F_n(t, \omega)$ is still uniformly convergent for almost all ω, the limit being $F(t, \omega \mid \mathcal{I})$, the conditional distribution function of X_1 given $\mathcal{I}$, the invariant field of the sequence. Another generalization of the Glivenko-Cantelli theorem was accomplished by Fisz [2], who noted that for each fixed n, $F_n(t, \omega)$ could be looked upon as a non-decreasing stochastic process and for each fixed t the sequence of arithmetic means derived from a sequence of independent random variables.

Looking at Tucker's theorem in this light, we could rephrase it as follows. Let X be a random variable, let $X(t) = I_{\{X \leqq t\}}$, and let T be a measure preserving set transformation. Choose $X_k(t) = T^k(X(t))$ in such a way that for each k, $X_k(t)$ is non-decreasing and right continuous. Then for almost all ω, $n^{-1}\sum_{k=1}^{n} X_k(t, \omega)$ converges uniformly in t. It is our purpose in this paper to show that this result remains true whenever $X(t)$ is any non-decreasing, right continuous process with $E(X(t))$ bounded. The proof is based on a general criterion for uniform convergence of a sequence of monotone processes and some results on conditional expectations which may prove of interest in themselves.

2. Conditional expectations for non-decreasing right continuous processes. Let $X(t, \omega)$ be a non-decreasing, right continuous process where t ranges over all real numbers. Let Y be an extended real valued function defined on our probability space Ω. We define

$$X(Y)(\omega) = X(Y(\omega), \omega).$$

For each real a, we define

$$Y_a(\omega) = \inf\{t : X(t, \omega) \geqq a\}.$$

The following result is then easily seen.

THEOREM 1. *For each real a and t, $\{Y_a > t\} = \{X(t) < a\}$ and $\{X(Y) < a\} = \{Y < Y_a\}$.*

In view of Theorem 1 it is obvious that Y_a is always $\mathfrak{F}(X(t): -\infty < t < \infty)$-measurable, and that $X(Y)$ is always measurable whenever Y is measurable.

Received 23 March 1965.

[1] This paper is part of a Ph.D. dissertation written at the University of Missouri under the direction of H. D. Brunk.

Now, let $\mathcal{J}$ be a fixed Borel field of events. It is our purpose in this section to find a workable representation of $E(X(Y) \mid \mathcal{J})$ when Y is $\mathcal{J}$-measurable.

We shall call $F(x, t, \omega \mid \mathcal{J})$ a conditional distribution function of the process $\{X(t)\}$ given $\mathcal{J}$ if for each fixed t it is a left continuous conditional distribution function of $X(t)$ given $\mathcal{J}$ and for each fixed ω and x a right continuous function of t. We shall call $E(t, \omega \mid \mathcal{J})$ a conditional expectation function of the process $\{X(t)\}$ if for each t, $E(t, \omega \mid \mathcal{J}) = E(X(t) \mid \mathcal{J})(\omega)$ a.s., and for each fixed ω, $E(t, \omega \mid \mathcal{J})$ is a right continuous non-decreasing function of t. The existence of such functions is a straightforward exercise.

LEMMA. *Let X and Y be extended real measurable functions with Y $\mathcal{J}$-measurable, and let $G(x, \omega \mid \mathcal{J})$ be a right continuous conditional distribution function of X given $\mathcal{J}$. For each ω in Ω let μ_ω be the measure on R with distribution function $G(x, \omega \mid \mathcal{J})$ and for each M in $R \times R$ let $M_\omega = \{x: (x, Y(\omega)) \varepsilon M\}$. Then $P((X,Y) \varepsilon M \mid \mathcal{J})(\omega) = \mu_\omega(M_\omega)$* a.s. *for each Borel set M in the finite plane.*

PROOF. We may clearly assume that $M = K \times L$ where K and L are one dimensional Borel sets. Then $\{(X, Y) \varepsilon M\} = \{X \varepsilon K\} \cap \{Y \varepsilon L\}$. Thus

$$P((X, Y) \varepsilon M \mid \mathcal{J})(\omega) = I_{\{Y \varepsilon L\}}(\omega) P(X \varepsilon K \mid \mathcal{J})(\omega) = I_{\{Y \varepsilon L\}}(\omega) \cdot \mu_\omega(K) = \mu_\omega(M_\omega).$$

THEOREM 2. *If Y is an $\mathcal{J}$-measurable random variable, then $E(\varphi(X(Y)) \mid \mathcal{J})(\omega) = \int \varphi(x)\, dF(x, Y(\omega), \omega \mid \mathcal{J})$* a.s., *whenever φ is a Borel function with $\varphi(X(Y))$ integrable.*

PROOF. It is clearly sufficient to prove the result when $\varphi = I_{(-\infty,a)}$ for a fixed real a. We then have

$$\begin{aligned} E(\varphi(X(Y) \mid \mathcal{J}) &= P(X(Y) < a \mid \mathcal{J}) = P(Y < Y_a \mid \mathcal{J}) \\ &= P(\{Y < Y_a\} \cap \{Y_a \neq \infty\} \mid \mathcal{J}) + P(Y_a = \infty \mid \mathcal{J}). \end{aligned}$$

Let $F_a(x, \omega \mid \mathcal{J})$ be a right continuous conditional distribution function of Y_a given $\mathcal{J}$. Then applying the lemma, with $G = F_a$ and $M = \{(x, y) \varepsilon R \times R: y < x\}$, we see that

$$\begin{aligned} E(\varphi(X(Y)) \mid \mathcal{J})(\omega) &= (F_a(\infty, \omega \mid \mathcal{J}) - F_a(Y(\omega), \omega \mid \mathcal{J}) + (1 - F_a(\infty, \omega \mid \mathcal{J})) \\ &= 1 - F_a(Y(\omega), \omega \mid \mathcal{J}) \quad \text{for all } \omega \end{aligned}$$

in an a.s. event A_1.

Now, for each fixed t, $1 - F_a(t, \omega \mid \mathcal{J}) = P(Y_a > t \mid \mathcal{J})(\omega) = P(X(t) < a \mid \mathcal{J}) = F(a, t, \omega \mid \mathcal{J})$ a.s. Thus there exists an almost sure event A_2 such that if ω is in A_2, $1 - F_a(t, \omega \mid \mathcal{J}) = F(a, t, \omega \mid \mathcal{J})$ for all rational t, and so, both functions being right continuous, for all real t. Thus, if ω is in $A_1 \cap A_2$, then

$$\begin{aligned} E(\varphi(X(Y) \mid \mathcal{J})(\omega) &= F(a, Y(\omega), \omega \mid \mathcal{J}) \\ &= \int \varphi(x)\, dF(x, Y(\omega), \omega \mid \mathcal{J}). \end{aligned}$$

THEOREM 3. *There exists an* a.s. *event A such that if ω is in A then*

$$\int x\, dF(x, t, \omega \mid \mathcal{J}) = E(t, \omega \mid \mathcal{J})$$

for all t.

PROOF. For each fixed t, $\int x\, dF(x, t, \omega \mid \mathcal{G}) = E(t, \omega \mid \mathcal{G})$ a.s. Also, if $t_1 < t_2$ and n is a non-negative integer then

$$\int_n^\infty x\, dF(x, t_1, \omega \mid \mathcal{G}) \leqq \int_n^\infty x\, dF(x, t_2, \omega \mid \mathcal{G}) \text{ a.s.}$$

For if $\alpha(x) = x \cdot I_{(n,\infty)}(x)$ then, α being non-decreasing, $\alpha(X(t_1)) \leqq \alpha(X(t_2))$ and

$$E(\alpha(X(t)) \mid \mathcal{G})(\omega) = \int \alpha(x)\, dF(x, t, \omega \mid \mathcal{G}) \text{ a.s.}$$

for each t. Applying a similar argument to integration from $-\infty$ to $-n$, it is easily seen that there is an almost sure event A such that if ω is in A, then

(1) $\int x\, dF(x, t, \omega \mid \mathcal{G}) = E(t, \omega \mid \mathcal{G})$ for all rational t,

(2) $\int_n^\infty x\, dF(x, t_1, \omega \mid \mathcal{G}) \leqq \int_n^\infty x\, dF(x, t_2, \omega \mid \mathcal{G})$ for all non-negative integers n and all rational $t_1 < t_2$,

(3) $\int_{-\infty}^{-n} x\, dF(x, t_1, \omega \mid \mathcal{G}) \leqq \int_{-\infty}^{-n} x\, dF(x, t_2, \omega \mid \mathcal{G})$ for all non-negative integers n and all rational $t_1 < t_2$.

Let ω be in A, let t be real and let $\{t_m\}$ be a decreasing sequence of rationals converging to t. The proof will be complete if we can show

$$\int x\, dF(x, t, \omega \mid \mathcal{G}) = \lim \int x\, dF(x, t_m, \omega \mid \mathcal{G}).$$

To see this let

$$\varphi_n(x) = -n \cdot I_{(-\infty,n)}(x) + x \cdot I_{[-n,n]} + n \cdot I_{(n,\infty)}(x).$$

Also let $a_{mn} = \int_0^\infty \varphi_n(x)\, dF(x, t_m, \omega \mid \mathcal{G})$. Then $\int_0^\infty \varphi_n(x)\, dF(x, t, \omega \mid \mathcal{G}) = \lim_{m=\infty} a_{mn}$, for, $\varphi_n \cdot I_{[0,\infty]}$ being bounded and continuous, the Helly Bray theorem applies ([3], p. 180). Thus

$$\int_0^\infty x\, dF(x, t, \omega \mid \mathcal{G}) = \lim_{n=\infty} \lim_{m=\infty} a_{mn}.$$

On the other hand

$$\begin{aligned} 0 &\leqq \int_0^\infty x\, dF(x, t_m, \omega \mid \mathcal{G}) - \int_0^\infty \varphi_n(x)\, dF(x, t_m, \omega \mid \mathcal{G}) \\ &= \int_n^\infty (x - n)\, dF(x, t_m, \omega \mid \mathcal{G}) \\ &\leqq \int_n^\infty x\, dF(x, t_m, \omega \mid \mathcal{G}) \\ &\leqq \int_n^\infty x\, dF(x, t_1, \omega \mid \mathcal{G}) \qquad \text{for all } m. \end{aligned}$$

It follows that a_{mn} converges to $\int_0^\infty x\, dF(x, t_m, \omega \mid \mathcal{G})$ uniformly in m. Also $\lim_{n=\infty} a_{mn}$ is a non-increasing sequence, bounded from below by 0. Thus

$$\int_0^\infty x\, dF(x, t, \omega \mid \mathcal{G}) = \lim_{m=\infty} \lim_{n=\infty} a_{mn} = \lim_{m=\infty} \int_0^\infty x\, dF(x, t_m, \omega \mid \mathcal{G}).$$

Now let $b_{mn} = \int_{-\infty}^0 \varphi_n(x)\, dF(x, t_m, \omega \mid \mathcal{G})$. By the same argument we see that

$$\int_{-\infty}^0 x\, dF(x, t, \omega \mid \mathcal{G}) = \lim_{n=\infty} \lim_{m=\infty} b_{mn}.$$

Also

$$
\begin{aligned}
0 &\leqq \textstyle\int_{-\infty}^{0} \varphi_n(x)\, dF(x, t_m, \omega \mid \mathcal{G}) - \int_{-\infty}^{0} x\, dF(x, t_m, \omega \mid \mathcal{G}) \\
&= -\textstyle\int_{-\infty}^{-n} (n + x)\, dF(x, t_m, \omega \mid \mathcal{G}) \\
&\leqq -\textstyle\int_{-\infty}^{-n} x\, dF(x, t_m, \omega \mid \mathcal{G}) \\
&\leqq -\textstyle\int_{-\infty}^{-n} x\, dF(x, t_0, \omega \mid \mathcal{G})
\end{aligned}
$$

where t_0 is a rational with $t_0 < t$. Thus b_{mn} converges to $\int_{-\infty}^{0} x\, dF(x, t_m, \omega \mid \mathcal{G})$ uniformly in m and $\lim_{n=\infty} b_{mn}$ is a non-increasing sequence bounded from below by $\int x\, dF(x, t_0, \omega \mid \mathcal{G})$. We then have

$$\int_{-\infty}^{0} x\, dF(x, t, \omega \mid \mathcal{G}) = \lim_{m=\infty} \lim_{n=\infty} b_{mn} = \lim_{m=\infty} \int_{-\infty}^{0} x\, dF(x, t_m, \omega \mid \mathcal{G})$$

and the desired result follows.

THEOREM 4. *Let Y be an $\mathcal{G}$-measurable random variable with $X(Y)$ integrable. Then* $E(X(Y)) \mid \mathcal{G})(\omega) = E(Y(\omega), \omega \mid \mathcal{G})$ a.s.

PROOF. This follows immediately from Theorem 2, with $\varphi(x) = x$, and Theorem 3.

It can be shown that if $E(X(t))$ is bounded from below then Theorem 4 remains true when Y is any extended real valued $\mathcal{G}$-measurable function which never takes the value $+\infty$.

3. A general theorem on uniform convergence. In this section we shall establish a criterion for the a.s. uniform on compacta convergence of a sequence of monotone processes. In the case where the probability space contains just one point a stronger form of our theorem is known from real variables: let F and each F_n be non-decreasing real functions defined on R. Suppose that $\lim F_n(t) = F(t)$ for all t in some dense subset and that, when t is a discontinuity point of F, then $\lim F_n(t) = F(t)$, $\lim F_n(t + 0) = F(t + 0)$ and $\lim F_n(t - 0) = F(t - 0)$. Then $\lim F_n(t) = F(t)$ uniformly on each compact interval of R. (The author is indebted to the referee for suggesting the use of this result, which greatly simplifies the proof of Theorem 5.)

THEOREM 5. *Let $\{X(t)\}$ and each $\{X_n(t)\}$ be non-decreasing processes, and let $\mathfrak{F}$ be the Borel field determined by $\{X(t)\colon -\infty < t < \infty\}$. Then there exists an* a.s. *A such that for all ω in A, $\lim X_n(t, \omega) = X(t, \omega)$ uniformly on each compact interval if and only if for each $\mathfrak{F}$-measurable random variable Y,*

$$\lim X_n(Y) = X(Y) \text{ a.s.,}$$

$$\lim X_n(Y + 0) = X(Y + 0) \text{ a.s.,}$$

and

$$\lim X_n(Y - 0) = X(Y - 0) \text{ a.s.}$$

PROOF. The necessity is valid for any real function Y, since uniformity of convergence allows the interchange of order in passing to the limit.

We now prove the sufficiency. For each real a, let $\bar{Y}_a(\omega) = \inf \{t\colon X(t + 0, \omega) \geqq a\}$ and then set

$$Y_a = \bar{Y}_a \cdot I_{\{|\bar{Y}_a| \neq \infty\}}.$$

By Theorem 1, Y_a is measurable with respect to the field determined by $\{X(t + 0: -\infty < t < \infty\}$ and is therefore $\mathfrak{F}$-measurable. Since constant functions are $\mathfrak{F}$-measurable, it is clear that there exists an a.s. event A such that if ω is in A then

$$\begin{aligned} \lim X_n(t, \omega) &= X(t, \omega) \quad \text{for all rational } t, \\ \lim X_n(Y_a(\omega), \omega) &= X(Y_a(\omega), \omega), \\ \lim X_n(Y_a(\omega) + 0, \omega) &= X(Y_a(\omega) + 0, \omega), \\ \lim X_n(Y_a(\omega) - 0, \omega) &= X(Y_a(\omega) - 0, \omega) \quad \text{for all rational } a. \end{aligned}$$

Let ω be in A. Suppose t_0 is a discontinuity point of $X(t, \omega)$. Pick a rational a, with

$$X(t_0 - 0, \omega) < a < X(t_0 + 0, \omega).$$

Then $\bar{Y}_a(\omega) \leqq t_0$. If $t < t_0$, then

$$X(t + 0, \omega) \leqq X(t_0 - 0, \omega) < a,$$

and so $\bar{Y}_a(\omega) > t$. Thus $t_0 = \bar{Y}_a(\omega) = Y_a(\omega)$. If we now set $F_n(t) = X_n(t, \omega)$ and $F(t) = X(t, \omega)$ then the conditions of the result mentioned at the beginning of this section are satisfied and the proof is complete.

We note that in general $X_n(Y)$ may not be measurable and so it appears that the theorem will not be useful unless the processes are either right or left continuous.

4. A uniform ergodic theorem. In this section we shall use the results and notation of [1]. $\{X(t)\}$ will again be a non-decreasing right continuous process and T a measure preserving set transformation defined on some Borel field containing $\mathfrak{F}(X(t): -\infty < t < \infty)$. We shall denote by $\mathfrak{I}$ the invariant field of T. It is easily seen that we may choose $X^*(t)$ in such a way that $X^*(t) = T(X(t))$ a.s. for all t and the process $\{X^*(t)\}$ is itself non-decreasing and right continuous. Moreover, if Y is an $\mathfrak{I}$-measurable random variable, then $X^*(Y) = T(X(Y))$ a.s. For let a be real. Let $Y_a^*(\omega) = \inf \{t: X^*(t, \omega) \geqq a\}$. Then

$$\begin{aligned} T(X(Y) < a) &= T(Y < Y_a) \\ &= \bigcup (T(Y < t) \cap T(Y_a > t): t \text{ is rational}) \\ &= \bigcup (\{Y < t\} \cap T(X(t) < a): t \text{ is rational}) \\ &= \bigcup (\{Y < t\} \cap \{X^*(t) < a\}: t \text{ is rational}) \\ &= \bigcup (\{Y < t\} \cap \{Y_a^* > t\}): t \text{ is rational}) \\ &= \{Y < Y_a^*\} = \{X^*(Y) < a\} \text{ a.s.} \end{aligned}$$

Thus, if $\varphi = I_{(-\infty, a)}$, $T(\varphi(X(Y))) = (\varphi(X^*(Y)))$ a.s. This clearly remains true for all Borel functions φ and the desired result is obtained by taking $\varphi(x) = x$.

Suppose now we use T to generate a sequence of processes, i.e., let $X_k(t) = T^k(X(t))$ for each k and t and choose the representations in such a way that for each k, $\{X_k(t)\}$ is a non-decreasing right continuous process. We will show that if $E(X(t))$ is bounded then there is uniform convergence of $X_k(t)$ in arithmetic means.

THEOREM 6. *Let $\{X(t)\}$ be a non-decreasing right continuous process with each $X(t)$ integrable. Let T be a measure preserving set transformation with invariant field $\mathcal{J}$ and let $E(t, \omega \mid \mathcal{J})$ be a conditional expectation function of $\{X(t)\}$ given $\mathcal{J}$. Pick $X_k(t) = T^k(X(t))$ in such a way that for each k, $\{X_k(t)\}$ is non-decreasing and right continuous. Then for almost all ω, $\lim n^{-1}\sum_{k=1}^{n} X_k(t, \omega) = E(t, \omega \mid \mathcal{J})$ uniformly on each compact interval. If, moreover, $E(X(t))$ is bounded, then for almost all ω the convergence is uniform on R.*

PROOF. Let $S_n(t) = n^{-1}\sum_{k=1}^{n} X_k(t)$. Let J be a fixed positive integer. To prove the first part of our assertion it is sufficient to show that for almost all ω we have uniform convergence on the interval $[-J, J]$. In showing this there is clearly no loss of generality in assuming that $X(t) = X(-J)$ if $t \leqq -J$ and $X(t) = X(J)$ if $t \geqq J$.

Let the random variable Y be measurable with respect to the field of the process $\{E(t, \omega \mid \mathcal{J})\}$. Then Y is $\mathcal{J}$-measurable. Moreover, $X(Y)$ is integrable, for $X(-J) \leqq X(Y) \leqq X(J)$, and we have seen that $X_k(Y) = T^k(X(Y))$. It therefore follows from the individual ergodic theorem that $\lim S_n(Y) = E(X(Y) \mid \mathcal{J})$ a.s. Thus, by Theorem 4, $\lim S_n(Y(\omega), \omega) = E(Y(\omega), \omega \mid \mathcal{J})$ a.s. Now set $X_k^*(t, \omega) = X_k(t - 0, \omega)$ and $E^*(t, \omega \mid \mathcal{J}) = E(t - 0, \omega \mid \mathcal{J})$. We may then replace $X_k(t, \omega)$, $E(t, \omega \mid \mathcal{J})$ and Y by $-X_k^*(-t, \omega)$, $-E^*(-t, \omega \mid \mathcal{J})$ and $-Y$ respectively in the above argument. This yields $\lim S_n(Y(\omega) - 0, \omega) = E(Y(\omega) - 0, \omega \mid \mathcal{J})$ a.s. We have shown that the conditions of Theorem 5 are satisfied and so the proof of the first half of the theorem is complete.

Now suppose $E(X(t))$ is bounded. Then $X(\pm\infty)$ are both a.s. finite and integrable. Also $X_k(\pm\infty) = T^k(X(\pm\infty))$ and $E(X(\pm\infty) \mid \mathcal{J})(\omega) = E(\pm\infty, \omega \mid \mathcal{J})$ a.s. Thus $\lim S_n(\pm\infty, \omega) = E(\pm\infty, \omega \mid \mathcal{J})$ and $E(\pm\infty, \omega \mid \mathcal{J})$ are finite for almost all ω. If ω is also a point at which $\lim S_n(t, \omega) = E(t, \omega \mid \mathcal{J})$ uniformly on each compact interval then, bearing in mind that the functions involved are non-decreasing, it is clear that the convergence is uniform on R.

REFERENCES

[1] DOOB, J. L. (1953). *Stochastic Processes*. Wiley, New York.

[2] FISZ, M. (1958). A limit theorem for non-decreasing random functions. *Bull. Acad. Polon. Sci. Ser. Sci. Math. Astronom. Phys.* **6** 485–497.

[3] LOÈVE, M. M. (1955). *Probability Theory*. Van Nostrand, Princeton.

[4] TUCKER, HOWARD G. (1959). A generalization of the Glivenko-Cantelli theorem. *Ann. Math. Statist.* **30** 828–830.