Proc. Funct. Anal. IV (Dubrovnik 1993), *Various Publ. Ser.* Vol. 43, 1994, (165-190)
Preprint Ser. No. 27, 1992, *Math. Inst. Aarhus*

Necessary and Sufficient Conditions for the Uniform Law of Large Numbers in the Stationary Case

GORAN PESKIR *and* MICHEL WEBER

Necessary and sufficient conditions for the uniform law of large numbers for stationary ergodic sequences of random variables are given. Three different types of conditions are investigated and established. Firstly it is shown that eventually total boundedness in mean is necessary and sufficient. This fact enables one to deduce the equivalence among almost sure convergence, convergence in mean, and convergence in probability in the uniform law under consideration. Secondly it is shown that eventual tightness is necessary and sufficient implying that the preceding convergence notions coincide with a weak convergence. Finally it is shown that the uniform convergence under consideration might be characterized in terms of the existence of certain totally bounded pseudo-metrics having a stochastic and mean control on increments of the process in question. Methods used in the proofs rely upon techniques established in the case of independent and identically distributed random variables interacting with those coming from ergodic theory. Given collaboration is significant and in this way some new facts are also obtained in the case of independent and identically distributed random variables itself.

1. Introduction

Let $\xi = \{ \xi_j \mid j \geq 1 \}$ be a sequence of independent identically distributed random variables defined on a probability space $(\Omega, \mathcal{F}, P)$ with values in a measurable space $(S, \mathcal{A})$ and a common distribution law π , let T be an arbitrary set, and let $f : S \times T \to \mathbf{R}$ be a given map. Suppose that the π-mean function M associated to f and given by:

$$M(t) = \int_S f(s,t)\ \pi(ds) \tag{1}$$

is well-defined for all $t \in T$, then in [10], [12], [28] and [32] one can find a series of necessary and sufficient conditions for the following form of the law of large numbers:

$$\frac{1}{n} \sum_{j=1}^{n} f(\xi_j, t)\ \rightrightarrows\ M(t) \tag{2}$$

AMS 1980 *subject classifications.* Primary 28D05, 60B12, 60F05, 60F15, 60F17, 60G10. Secondary 28A20.
Key words and phrases: Stationary, ergodic, uniform law of large numbers, eventually total boundedness in mean, eventual tightness, perfectness, uniform integrability. © goran@imf.au.dk

where a probabilistic convergence is uniform over $t \in T$. Under the expression "a probabilistic convergence" we mean either of the following convergence notions (or their non-measurable relatives): *almost sure convergence, convergence in mean, convergence in probability, weak convergence*. And roughly speaking it is shown in [12] that *eventually total boundedness in π-mean* of f is necessary and sufficient for (2). A martingale extension of these results is given in [23] by showing that *Hardy's regular convergence* of corresponding means is necessary and sufficient for a useful consequence of (2), see also [14]. Let us say that given results have found nice applications in establishing consistency of statistical models, see [16] and [22].

In this paper we consider a new case of the same problem of uniform convergence which is not covered by the preceding two ones. Similarly to the martingale approach we drop independence, but now instead of martingale property we suppose stationarity. In other words we shall assume that the sequence $\xi = \{ \xi_j \mid j \geq 1 \}$ under consideration is a *stationary* sequence of random variables. By the law of large numbers for stationary sequences we have:

$$\frac{1}{n} \sum_{j=1}^{n} f(\xi_j, t) \to L(t) \quad P\text{-}a.s. \tag{3}$$

as $n \to \infty$ for $t \in T$, but this time the limit $L(t)$ need not to be degenerated as it was in the case of independent and identically distributed random variables. However this happens if the stationary sequence ξ is *ergodic*, while in order to establish ergodicity of ξ in general, it is enough to have *Blackwell-Freedmann's asymptotic independence condition* saying that the tail σ-algebra associated to ξ is degenerated. This fact would provide a clear connection between ergodicity and dropping independence. More details in this direction will be presented later. Of course if the limit $L(t)$ in (2) is degenerated, then it coincides with the π-mean function $M(t)$ for $t \in T$ given by (1). Let us moreover say that our main motivation for present work is coming from statistics where the limit L plays a role of so-called information function, see [16], and a statistical background requires to assume that L is degenerated. Therefore we shall be mainly concerned on considerations taking this assumption into account. Others justifications for this direction are coming both from theory and applications. Namely it is a well-known result in the theory of stationary processes that any stationary process may be represented as a convex combination of stationary ergodic processes, see [4] (p.221). Moreover after a certain dose of experience with concrete examples one may easily conclude that modeling a system with a non-ergodic process is usually a wrong business. In other words we may say that the system under consideration is not completely decomposed into its separate component parts. Finally a long time ago Gibb's conjecture resulted with a well-known hypothesis in statistical mechanics on equality among the space and time average induced by endomorphisms arising in the Hamiltonian flow in a phase space. Von Neumann's mean and Birkhoff's pointwise theorem made this hypothesis clear for *ergodic* endomorphisms. And in order to account all these facts it is very natural to assume that the sequence $\xi = \{ \xi_j \mid j \geq 1 \}$ under our next considerations on the uniform convergence is both stationary and ergodic as well. Let us in addition emphasize that all informations on stationarity and ergodicity of ξ are contained in the distribution law P_ξ of ξ in the countable product $(S^{\mathbf{N}}, \mathcal{A}^{\mathbf{N}})$ of $(S, \mathcal{A})$ with itself. Therefore we may in next follow a former approach in [12] and suppose that our probability space $(\Omega, \mathcal{F}, P)$ equals $(S^{\mathbf{N}}, \mathcal{A}^{\mathbf{N}}, P_\xi)$. The sequence ξ is then given by $\xi_j(\omega) = \omega_j$ for $\omega = (\omega_i) \in S^{\mathbf{N}}$

and $j \geq 1$, and the given realization of ξ is usually called *the canonical representation* of ξ . In this way one can avoid certain measure-theoretic pathologies and make things convenient for a smooth work. However examples due to Gerstenhaber [11] (p.32), Tanny [30] and Aaronson [1] show that unavoidable problems may occur for the same conclusions in a general stationary ergodic case without independence or additional hypothesis. Therefore we shall in next mainly work with a general probability space $(\Omega, \mathcal{F}, P)$. It turns out that above slight restriction to the canonical representation may be removed by assuming the perfectness on ξ as a map from Ω into $S^{\mathbf{N}}$. To conclude this introduction let us say that the methods used in the proofs rely upon techniques established in the case of independent and identically distributed random variables in [12] interacting with those coming from ergodic theory. Given collaboration is significant and in this way some new facts are also obtained in the case of independent and identically distributed random variables itself.

2. Preliminary facts

We shall begin by recalling some definitions and results which will be of use in the rest of the paper or just increase the clarity of the exposed material. First we consider stationary sequences of random variables. Roughly speaking these sequences might be described as those whose distributions remain unchanged as time passes. More precisely, a sequence of random variables $\{ \xi_j \mid j \geq 1 \}$ is said to be *stationary*, if we have:

$$(1) \qquad (\xi_{n_1}, \dots, \xi_{n_k}) \sim (\xi_{n_1+\tau}, \dots, \xi_{n_k+\tau})$$

for all $1 \leq n_1 < \ldots < n_k$ and all $\tau = 1, 2 \ldots$. In order to shed some more light upon stationary sequences, we shall recall some notions from ergodic theory. Let $(\Omega, \mathcal{F}, P)$ be a probability space, then a map τ from Ω into itself is said to be *an endomorphism*, if it is measurable and satisfies $P \circ \tau^{-1} = P$. Endomorphisms are often called *measure preserving transformations*. Let τ be an endomorphism in $(\Omega, \mathcal{F}, P)$, then a set A in $\mathcal{F}$ is said to be τ*-invariant*, if $\tau^{-1}(A) = A$. The family $\mathcal{F}_\tau$ of all τ-invariant sets in $\mathcal{F}$ is a σ-algebra in Ω . An endomorphism τ is called *ergodic*, if $P(A) \in \{0,1\}$ for all $A \in \mathcal{F}_\tau$. A measurable function f from Ω into $\mathbf{R}$ is called τ*-invariant*, if $f \circ \tau = f$. It is easily verified that f is τ-invariant, if and only if it is $\mathcal{F}_\tau$-measurable. Therefore for ergodic τ any τ-invariant function almost surely equals to a constant. An endomorphism τ in $(\Omega, \mathcal{F}, P)$ is called *mixing*, if we have:

$$\lim_{i \to \infty} P(A \cap \tau^{-i}(B)) = P(A) \cdot P(B)$$

for all $A, B \in \mathcal{F}$. If τ is mixing, then it is obviously ergodic. Let us now consider a measurable space $(S, \mathcal{A})$, and let $(S^{\mathbf{N}}, \mathcal{A}^{\mathbf{N}})$ denote the countable product of $(S, \mathcal{A})$ with itself. Then *the unilateral shift* θ is a map from $S^{\mathbf{N}}$ into $S^{\mathbf{N}}$ defined by:

$$\theta(s_1, s_2, s_3 \dots) = (s_2, s_3 \dots)$$

for all $(s_1, s_2, \dots) \in S^{\mathbf{N}}$. Let π be a probability measure on $(S, \mathcal{A})$, and let $\pi^{\mathbf{N}}$ be the countable product of π with itself. Then θ is an endomorphism in $(S^{\mathbf{N}}, \mathcal{A}^{\mathbf{N}}, \pi^{\mathbf{N}})$ and it is well-known that θ is mixing and thus ergodic. Therefore Kolmogorov's strong law of large numbers follows straight forward by Birkhoff's pointwise ergodic theorem, see [17] (p.24). We are now in position to define stationarity in an instructive way. Let $\xi = \{ \, \xi_j \mid j \geq 1 \, \}$ be a sequence of random variables defined on a probability space $(\Omega, \mathcal{F}, P)$ with values in a measurable space $(S, \mathcal{A})$, and let P_ξ be the distribution law of ξ as a random variable from Ω into $S^{\mathbf{N}}$, that is:

$$P_\xi(A) = P\{\xi \in A\}$$

for all $A \in \mathcal{A}^{\mathbf{N}}$. Then P_ξ is a probability measure on $(S^{\mathbf{N}}, \mathcal{A}^{\mathbf{N}})$, and ξ is said to be *stationary*, if the unilateral shift θ is a measure preserving transformation in $(S^{\mathbf{N}}, \mathcal{A}^{\mathbf{N}}, P_\xi)$. It is quite clear that the present definition coincides with that given by (1). And a stationary sequence ξ is called *ergodic*, if the unilateral shift θ is ergodic in $(S^{\mathbf{N}}, \mathcal{A}^{\mathbf{N}}, P_\xi)$. The next four statements will provide useful facts on stationary and ergodic sequences of random variables.

(2) *(Endomorphisms generate plenty stationary sequences)*

Let τ be an endomorphism in a probability space $(\Omega, \mathcal{F}, P)$, let $(S, \mathcal{A})$ be a measurable space, and let f be a measurable map from Ω into S. If we define $\xi_{i+1} = f \circ \tau^i$ for $i = 0, 1 \ldots$, then $\xi = \{ \, \xi_j \mid j \geq 1 \, \}$ is a stationary sequence of random variables from Ω into S. Moreover if τ is ergodic, then ξ is also ergodic.

(3) *(Shift keep stationarity)*

Let $\xi = \{ \, \xi_j \mid j \geq 1 \, \}$ be a stationary sequence of random variables defined on a probability space $(\Omega, \mathcal{F}, P)$ with values in a measurable space $(S, \mathcal{A})$, let θ be the unilateral shift in $S^{\mathbf{N}}$, let $(T, \mathcal{B})$ be an another measurable space, and let $F : S^{\mathbf{N}} \rightarrow T$ be a measurable map. If we define:

$$\eta_{i+1} = F \circ \theta^i \circ \xi$$

for $i = 0, 1 \ldots$, then $\eta = \{ \, \eta_j \mid j \geq 1 \, \}$ is a stationary sequence of random variables from Ω into T. In particular if $\xi = \{ \, \xi_j \mid j \geq 1 \, \}$ is a stationary sequence of real valued random variables, and if we consider:

$$\sigma_i = \frac{1}{N} \sum_{j=i}^{i+N-1} \xi_j$$

for given and fixed $N \geq 1$ and all $i \geq 1$, then $\sigma = \{ \, \sigma_j \mid j \geq 1 \, \}$ is stationary.

(4) *(Shift keep ergodicity)*

Under the assumptions in (3) suppose moreover that the sequence $\xi = \{ \, \xi_j \mid j \geq 1 \, \}$ is ergodic, then $\eta = \{ \, \eta_j \mid j \geq 1 \, \}$ is also ergodic.

(5) *(The law of large numbers for stationary sequences)*

Let $\xi = \{ \, \xi_j \mid j \geq 1 \, \}$ be a stationary sequence of real valued random variables defined on a probability space $(\Omega, \mathcal{F}, P)$, let θ be the unilateral shift in $\mathbf{R}^{\mathbf{N}}$, and let $\mathcal{B}_\theta(\mathbf{R}^{\mathbf{N}})$ denote the σ-algebra of all θ-invariant sets in the Borel σ-algebra $\mathcal{B}(\mathbf{R}^{\mathbf{N}})$. If ξ_1 belongs *to* $L^1(P)$ *, then we have:*

$$\frac{1}{n}\sum_{j=1}^{n}\xi_j \;\rightarrow\; E\{\,\xi_1 \mid \xi^{-1}(\mathcal{B}_\theta(\mathbf{R}^{\mathbf{N}}))\}$$

P*-a.s. and in* P*-mean, as* $n \rightarrow \infty$.

The proofs of (2)-(4) may be deduced straight forward by definition, while the proof of (5) may be easily established by using Birkhoff's pointwise ergodic theorem for almost sure convergence, and von Neumann's mean ergodic theorem for convergence in mean. Let us emphasize that convergence in mean may be also obtained from Birkhoff's theorem by using the following trivial but useful result on uniform integrability:

(6) *Let* $\xi = \{\,\xi_j \mid j \geq 1\,\}$ *be a stationary sequence of real valued random variables defined on a probability space* $(\Omega, \mathcal{F}, P)$ *such that* ξ_1 *belongs to* $L^1(P)$ *. Then the sequence of averages:*

$$A = \{\,\frac{1}{n}\sum_{j=1}^{n}\xi_j \;\mid\; n \geq 1\,\}$$

is uniformly integrable.

Let us now pass to ergodicity of stationary sequences. Preliminary facts in this direction are already given in introduction and we proceed with some more technical details. Let us first of all notice that if the sequence $\xi = \{\,\xi_j \mid j \geq 1\,\}$ in (5) is ergodic, then we clearly have:

$$\frac{1}{n}\sum_{j=1}^{n}\xi_j \;\rightarrow\; E(\xi_1) \tag{7}$$

as $n \rightarrow \infty$. In order to obtain a better understanding of what is going on under this condition, we shall distinguish two approaches. First we turn out some very general necessary and sufficient conditions for ergodicity itself, and then we look for a connection with independence by giving sufficient conditions. As an example we recall the Gaussian case where given conditions become of a particularly concrete form. Let $\xi = \{\,\xi_j \mid j \geq 1\,\}$ be a sequence of random variables defined on a probability space $(\Omega, \mathcal{F}, P)$ with values in a measurable space $(S, \mathcal{A})$, let P_ξ denote the distribution law of ξ in $(S^{\mathbf{N}}, \mathcal{A}^{\mathbf{N}})$, and let θ denote the unilateral shift in $(S^{\mathbf{N}}, \mathcal{A}^{\mathbf{N}})$.

(8) *Suppose that given* ξ *is stationary, then it is ergodic, if and only if either of the following four equivalent conditions is satisfied:*

(i) $P_\xi(A) > 0$ *for* $A \in \mathcal{A}^{\mathbf{N}} \;\Rightarrow\; P_\xi(\bigcup_{i=0}^{\infty}\theta^{-i}(A)) = 1$

(ii) $P_\xi(A)\cdot P_\xi(B) > 0$ *for* $A, B \in \mathcal{A}^{\mathbf{N}} \;\Rightarrow\; \sum_{i=0}^{\infty} P_\xi(A \cap \theta^{-i}(B)) > 0$

(iii) $n^{-1} \sum_{j=0}^{n-1} P_\xi(A \cap \theta^{-j}(B)) \to P_\xi(A) \cdot P_\xi(B)$ *, for all* $A, B \in \mathcal{A}^{\mathbf{N}}$

(iv) $n^{-1} \sum_{j=0}^{n-1} \int_{S^{\mathbf{N}}} F \cdot (G \circ \theta^j)\, dP_\xi \to \int_{S^{\mathbf{N}}} F\, dP_\xi \cdot \int_{S^{\mathbf{N}}} G\, dP_\xi$ *, for all* $F, G \in L^2(P_\xi)$.

We proceed by considering a connection between ergodicity and independence. Let p_i denote the i-th projection from $S^{\mathbf{N}}$ into S for $i \geq 1$. Then *the tail σ-algebra* in $(S^{\mathbf{N}}, \mathcal{A}^{\mathbf{N}})$ is defined by $\mathcal{A}_\infty = \bigcap_{n=1}^\infty \sigma(p_j \mid j \geq n)$, and we obviously have:

$$\mathcal{A}_\theta \subset \mathcal{A}_\infty \subset \mathcal{A}^{\mathbf{N}}$$

where $\mathcal{A}_\theta$ denotes as usual the σ-algebra of all θ-invariant sets in $\mathcal{A}^{\mathbf{N}}$. Therefore the triviality of the tail σ-algebra $\mathcal{A}_\infty$ will imply the ergodicity of the unilateral shift in $(S^{\mathbf{N}}, \mathcal{A}^{\mathbf{N}}, P_\xi)$. It is well-known that this happens, if and only if the following *Blackwell-Freedmann asymptotic independence condition* is satisfied, see [17] (p.28):

$$\lim_{n \to \infty} \sup_{B \in \sigma(p_j | j \geq n)} |\, P_\xi(A \cap B) - P_\xi(A) \cdot P_\xi(B)\, | = 0 \tag{9}$$

for every $A \in \mathcal{A}_\infty$. Let us in addition note that random variables from a stationary sequence are identically distributed. Of course every sequence of *independent and identically distributed* random variables is stationary and ergodic. So-called *moving averages* are also known to be stationary and ergodic. Others important examples of stationary ergodic sequences may be found in the theory of *Markov chains*. Notions like *almost periodic functions*, *white noise*, *autoregression* and *balance equations* are also of vital importance in this direction, as well as various *mixed models*. For more informations in this direction we shall refer the reader to [6] and [27]. But let us say that in the study of stationary processes in general, one class takes a central place. These are the *Gaussian* stationary processes. Let $\xi = \{\, \xi_j \mid j \geq 1 \,\}$ be a Gaussian sequence with *mean* 0 and *covariance* Γ , that is $E(\xi_i) = 0$ and $\Gamma(i,j) = E(\xi_i \xi_j)$ for $i, j \geq 1$. Then ξ is stationary, if and only if $\Gamma(i,j) = R(i-j)$ for $i, j \geq 1$. The function R is called *the covariance function* of ξ . By Herglotz's theorem there exists a finite measure σ on the Borel σ-algebra in $[-\pi, \pi[$ such that:

$$R(n) = \int_{-\pi}^{\pi} \cos(n \cdot \lambda)\; \sigma(d\lambda) \tag{10}$$

for all $n \geq 0$, see [27] (p.393). Given measure σ is called *the spectral masure* of ξ . And it is well-known that ξ is ergodic, if and only if the spectral measure σ is continuous. This happens for instance if $R(n) \to 0$ as $n \to \infty$. Relation (10) is called *the spectral representation* of the covariance function R . A similar spectral representation of the sequence ξ itself is possible. In this way so-called *orthogonal stochastic measures* appear naturally, and an appropriate stochastic integration should be developed. Given representations generalize to complex valued stationary processes *in the wide sense* as well, for more details see [27]. To conclude these considerations we recall a *zero-one law for stationary ergodic sequences* due to Tanny in [30]:

(11) *Let* $\xi = \{\, \xi_j \mid j \geq 1 \,\}$ *be a stationary ergodic sequence of real valued random variables, then for every decreasing sequence of Borel sets* $\{\, B_j \mid j \geq 1 \,\}$ *on the real line, the probability of* $A = \limsup_{i \to \infty} \{\xi_i \in B_i\}$ *equals* 0 *or* 1 .

The original proof of this fact is rather complicated. However we may present a simple proof as follows. First notice that there is no restriction to consider the canonical representation of ξ , see section 1, and let θ denote the unilateral shift in $\mathbf{R}^{\mathbf{N}}$. Since given B_j's are decreasing, then it is easily established that $\theta(A) \subset A$. Therefore $A \subset \theta^{-j}(A)$ for all $j = 1, 2 \ \dots$, and (11) follows straight forward by applying (iii) in (8) with $B = A$.

We proceed by recalling some facts from the calculus of non-measurable sets and functions. Let $(\Omega, \mathcal{F}, P)$ be a probability space, then P^* and P_* denotes the outer and inner P-measure. An arbitrary map ζ from Ω into $\bar{\mathbf{R}}$ is called *a random element*. If ζ is a random element on Ω , then $\int^* \zeta \ dP$ and $\int_* \zeta \ dP$ denotes the upper and lower P-integral of ζ , and ζ^* and ζ_* denotes the upper and lower P-envelope of ζ . Let $\{ \ \zeta_n \mid n \geq 1 \ \}$ be a sequence of real valued random elements defined on $(\Omega, \mathcal{F}, P)$. In next we shall follow a former approach presented in [23] and use the following convergence notions:

(a) $\zeta_n \rightarrow 0 \ \ (a.s.)$, if $\exists N \in \mathcal{F}$ such that $P(N) = 0$ and $\zeta_n(\omega) \rightarrow 0$, $\forall \omega \in \Omega \setminus N$

(b) $\zeta_n \rightarrow 0 \ \ (a.s.)^*$, if $|\zeta_n|^* \rightarrow 0 \ \ (a.s.)$

(c) $\zeta_n \rightarrow 0 \ \ (P^*)$, if $P^*\{ \ |\zeta_n| \geq \epsilon \ \} \rightarrow 0 \ , \ \forall \epsilon > 0$

(d) $\zeta_n \rightarrow 0 \ \ (P_*)$, if $P_*\{ \ |\zeta_n| \geq \epsilon \ \} \rightarrow 0 \ , \ \forall \epsilon > 0$

(e) $\zeta_n \rightarrow 0 \ \ (L^1)^*$, if $\int^* |\zeta_n| \ dP \rightarrow 0$

(f) $\zeta_n \rightarrow 0 \ \ (L^1)_*$, if $\int_* |\zeta_n| \ dP \rightarrow 0$.

Note if ζ_n is measurable for $n \geq 1$, then the convergence notion in (a) and (b) coincides with the notion of *P-almost sure convergence*, the convergence notion in (c) and (d) coincides with the notion of *convergence in P-probability*, and the convergence notion in (e) and (f) coincides with the notion of *convergence in P-mean*. For more details in this direction we shall refer the reader to [20], [21] and [23]. It is easily established that we have:

(12)
$$\begin{tikzcd}
(a.s.)^* \arrow[r, Rightarrow] \arrow[d, Rightarrow] & (a.s.) \arrow[d, Rightarrow] \\
(P^*) \arrow[r, Rightarrow] & (P_*) \\
(L^1)^* \arrow[r, Rightarrow] \arrow[u, Rightarrow] & (L^1)_* \arrow[u, Rightarrow]
\end{tikzcd}$$

and no other implication holds in general. Let ξ be a random variable defined on a probability space $(\Omega, \mathcal{F}, P)$ with values in a measurable space $(S, \mathcal{A})$. Then ξ is said to be *P-perfect*, if $\forall F \in \mathcal{F}$, $\exists A \in \mathcal{A}$, $A \subset \xi(F)$ such that $P(F \setminus \xi^{-1}(A)) = 0$. If ξ is P-perfect, then for any function Ψ from S into $\bar{\mathbf{R}}$ we have $\int^* \Psi \circ \xi \ dP = \int^* \Psi \ dP_\xi$ and $(\Psi \circ \xi)^* = \Psi^* \circ \xi$. For more informations in this direction see [20]. Let $(\Omega, \mathcal{F}, P)$ be a probability space, then we define $L^{<1>}(P)$ to be the set of all functions ζ from Ω into $\mathbf{R}$ satisfying $\| \ \zeta \ \|_1^* = \int^* | \ \zeta \ | \ dP < \infty$. Then $(L^{<1>}(P), \| \cdot \|_1^*)$ is a Banach space, and for more details see [21]. Let us in addition clarify if T is a non-empty set, then $\mathbf{R}^T$ denotes the set of all real valued functions defined on T , and $B(T)$ denotes the set of all *bounded* functions in $\mathbf{R}^T$. For $f \in \mathbf{R}^T$ and $A \subset T$ we put $\| f \|_A = \sup_{t \in A} | f(t) |$. Then $(f,g) \mapsto \| f - g \|_T$ defines a metric on $\mathbf{R}^T$, not necessarily finite valued, but topologically equivalent to the bounded metric $(f,g) \mapsto \arctan \| f - g \|_T$. It is well-known that $(B(T), \| \cdot \|_T)$ is a Banach space. A finite *cover* of a set T is a family $\gamma = \{ D_1, \ldots, D_n \}$ of non-empty subsets of T satisfying $T = \bigcup_{j=1}^n D_j$. The family of all finite covers of a set T will be denoted by $\Gamma(T)$.

3. The uniform law of large numbers in the stationary case

In this section we present basic results. First we formulate the problems under next consideration and discuss some necessary conditions. Let $\xi = \{ \xi_j \mid j \geq 1 \}$ be a stationary sequence of random variables defined on a probability space $(\Omega, \mathcal{F}, P)$ with values in a measurable space $(S, \mathcal{A})$ and a common distribution law π , let T be an arbitrary set, and let $f : S \times T \to \mathbf{R}$ be a given map. Let us for every $n \geq 1$ denote:

$$S_n(f) = \sum_{j=1}^{n} f(\xi_j) \ . \tag{1}$$

Then $S_n(f)$ maps Ω into $\mathbf{R}^T$, and we shall study the uniform convergence of the sequence of averages $A(f) = \{ n^{-1} S_n(f) \mid n \geq 1 \}$ over the set T . In other words we shall search necessary and sufficient conditions for:

$$\Big\| \frac{1}{n} S_n(f) - L \Big\|_T \to 0 \quad (c) \tag{2}$$

with some L in $\mathbf{R}^T$ as $n \to \infty$, where (c) denotes either of the following convergence notions: P-almost sure convergence, convergence in P-mean, convergence in P-probability. We shall also consider a weak convergence related to (2). However let us note that the map on the left side in (2) needs not to be P-measurable under our general assumptions on the set T . Thus in order to maintain this generality we need to use some results from the calculus of non-measurable sets and functions. According to the facts on convergence notions of random elements presented in the last section, see (2.12), we may clearly conclude that $(a.s.)^*$-convergence, $(L^1)^*$-convergence and (P^*)-convergence are convergence notions of vital interest in (2). Therefore we shall soon turn our attention in this direction. All needed facts for establishing a weak convergence related to (2) will be presented later. For those who would like to avoid measurability problems in (2) we suggest to forget all "stars" in the notation and to assume measurability everywhere where is needed. This approach may be justified in a quite general setting by using the projection theorem, see [16]. Let us now consider the following two additional conditions on the map f :

$$\int^* \| f(s) \|_T \ \pi(ds) < \infty \tag{3}$$

$$s \mapsto f(s,t) \ \text{ is } \pi\text{-measurable for every } \ t \in T \ . \tag{4}$$

Under the canonical representation in the case of independent and identically distributed random variables it is shown that (3) is necessary for $(a.s.)$-convergence in (2), see [12]. Moreover in this case ξ_1 is P-perfect and thus (3) is equivalent to the following condition:

$$\int^* \| f(\xi_1) \|_T \ dP < \infty \ . \tag{5}$$

However an example due to Aaronson in [1] shows that this condition may fail in the general stationary case without additional hypothesis. Given example is a modification of an example due to Tanny in [30] where some more facts in this direction may be found. Actually an example of that type is also constructed a long time ago by Gerstenhaber, see [11] (p.32). In this way we are put in position *to assume that our map* f *satisfies (3).* Of course this is a very weak assumption and the establishment of this result from (2) has mainly a theoretical importance. Note that condition (3) may be simply expressed by requiring that $\| f \|_T$ belongs to $L^{<1>}(\pi)$, as well as condition (5) by requiring that $\| f(\xi_1) \|_T$ belongs to $L^{<1>}(P)$. Also note that (5) becomes a consequence of (3), and it is equivalent to (3) whenever ξ_1 is P-perfect. Let us now pass to condition (4). Similarly to (3) under the canonical representation in the case of independent and identically distributed random variables it is shown that (4) is necessary for (P^*)-convergence in (2). Therefore it is not a big restriction *to assume that our map* f *satisfies (4)*, and we shall permanently follow this course. Let us in addition say that *our stationary sequence* ξ *in next is supposed to be ergodic*. Several good and deep reasons justifying this assumption are stated in the first two sections. Moreover in this case by (2.3) and (2.4) we may also conclude that the sequence $\{ f(\xi_j, t) \mid j \geq 1 \}$ is stationary and ergodic for every $t \in T$. Therefore by (2.5) we may deduce:

$$\frac{1}{n} \sum_{j=1}^{n} f(\xi_j, t) \to M(t) \quad P\text{-}a.s. \tag{6}$$

as $n \to \infty$, where $M(t) = \int_S f(s,t) \ \pi(ds)$ is the π-mean function of f for every $t \in T$. Thus *we may assume that* L *in (2) equals* M . Note that *(3) and (4) would imply that* $f(t)$ *belongs to* $L^1(\pi)$ *for every* $t \in T$. Moreover *by assuming (3) we may easily establish that* M *belongs to* $B(T)$. This conclusion ends the discussion on our basic hypothesis. Let us say that their necessity for (2) under eventual additional hypothesis will be not considered in this paper.

We now pass to the uniform convergence itself. For this we shall recall that given f is said to be *eventually totally bounded in* π*-mean*, if the following condition is satisfied, see [12]:

(7) $\forall \epsilon > 0 \ , \ \exists \gamma_\epsilon \in \Gamma(T)$ such that:

$$\inf_{n \geq 1} \ \frac{1}{n} \int^* \sup_{t',t'' \in A} \ | \ S_n\big(f(t') - f(t'')\big) \ | \ dP < \epsilon$$

for all $A \in \gamma_\epsilon$.

To be precise, conditions (3) and (4) should be also included into definition, but for our purposes the present definition is more convenient and we shall recall (3) and (4) separately when needed. And our next aim is to show that (7) is necessary and sufficient for (2) with any of above mentioned convergence notions. First result in this direction may be stated as follows.

Theorem 1.

Let $\xi = \{ \xi_j \mid j \geq 1 \}$ *be a stationary ergodic sequence of random variables defined on a probability space* $(\Omega, \mathcal{F}, P)$ *with values in a measurable space* $(S, \mathcal{A})$ *and a common distribution law* π *, let* T *be an arbitrary set, and let* $f : S \times T \rightarrow \mathbf{R}$ *be a given map. Let us suppose that* $\| f \|_T$ *belongs to* $L^{<1>}(\pi)$ *and that the map* $s \mapsto f(s,t)$ *is* π*-measurable for every* $t \in T$ *, and let* $M(t) = \int_S f(s,t) \; \pi(ds)$ *be the* π*-mean function of* f *for* $t \in T$ *. If* f *is eventually totally bounded in* π*-mean and* ξ *is* P*-perfect as a map from* Ω *into* $S^{\mathbf{N}}$ *, then:*

$$\| \frac{1}{n} \sum_{j=1}^{n} f(\xi_j) - M \|_T \rightarrow 0 \;\; (a.s.)^* \;\&\; (L^1)^* \tag{1}$$

as $n \rightarrow \infty$ *.*

Proof. Let $\epsilon > 0$ be given, then there exists $\gamma_\epsilon \in \Gamma(T)$ satisfying:

$$\inf_{n \geq 1} \; \frac{1}{n} \int^* \sup_{t',t'' \in A} \mid S_n\big(f(t') - f(t'')\big) \mid \; dP < \epsilon \tag{2}$$

for all $A \in \gamma_\epsilon$. Since under our hypothesis M belongs to $B(T)$, then it is not a restriction to assume that we also have:

$$\sup_{t',t'' \in A} \mid M(t') - M(t'') \mid \; \leq \; \epsilon \tag{3}$$

for all $A \in \gamma_\epsilon$. Now for every A in γ_ϵ choose a point t_A in A . Then by (3) we have:

$$\begin{aligned} \| \frac{1}{n} S_n(f) - M \|_T = \; & \max_{A \in \gamma_\epsilon} \| \frac{1}{n} S_n(f) - M \|_A \leq \max_{A \in \gamma_\epsilon} \Big\{ \sup_{t',t'' \in A} \mid \frac{1}{n} S_n(f(t')) \\ & - \frac{1}{n} S_n(f(t'')) \mid \; + \; \mid \frac{1}{n} S_n(f(t_A)) - M(t_A) \mid \\ & + \sup_{t',t'' \in A} \mid M(t') - M(t'') \mid \Big\} \; \leq \; \max_{A \in \gamma_\epsilon} \sup_{t',t'' \in A} \mid \frac{1}{n} S_n\big(f(t') - f(t'')\big) \mid \\ & + \max_{A \in \gamma_\epsilon} \mid \frac{1}{n} S_n(t_A)) - M(t_A) \mid \; + \; \epsilon \; . \end{aligned}$$

By the law of large numbers for stationary sequences, see (2.5), hence we may easily deduce:

$$\begin{aligned} \limsup_{n \rightarrow \infty} \; \| \frac{1}{n} S_n(f) - M \|_T^* \leq \; & \limsup_{n \rightarrow \infty} \; \max_{A \in \gamma_\epsilon} \Big(\sup_{t',t'' \in A} \mid \frac{1}{n} S_n\big(f(t') - f(t'')\big) \mid \Big)^* + \; \epsilon \\ = \; & \max_{A \in \gamma_\epsilon} \; \limsup_{n \rightarrow \infty} \Big(\sup_{t',t'' \in A} \mid \frac{1}{n} S_n\big(f(t') - f(t'')\big) \mid \Big)^* + \; \epsilon \; . \end{aligned}$$

Therefore by using (2) we may conclude that in order to establish $(a.s.)^*$-convergence in (1) it is enough to show the following inequality:

$$
\text{(4)} \qquad \limsup_{n\to\infty} \Big(\sup_{t',t''\in A} \mid \frac{1}{n} S_n\big(f(t')-f(t'')\big) \mid \Big)^* \leq
$$
$$
\leq \inf_{n\geq 1} \int^* \sup_{t',t''\in A} \mid \frac{1}{n} S_n\big(f(t')-f(t'')\big) \mid dP
$$

for every $A \in \gamma_\epsilon$. We leave this fact to be established with some more details in the next proposition. Thus we may proceed with $(L^1)^*$-convergence in (1). Let F denote the upper π-envelope of the map $s \mapsto \| f(s)-M \|_T$. By our assumptions we may easily verify that F belongs to $L^1(\pi)$. Therefore by (2.3) and (2.6) the sequence of averages $\{ n^{-1} \sum_{j=1}^n F(\xi_j) \mid n \geq 1 \}$ is uniformly integrable. Now note that we have:

$$
\| \frac{1}{n} S_n(f) - M \|_T^* \leq \frac{1}{n} \sum_{j=1}^n F(\xi_j)
$$

for all $n \geq 1$, and thus the sequence $\{ \| n^{-1} S_n(f) - M \|_T^* \mid n \geq 1 \}$ is uniformly integrable as well. Therefore $(L^1)^*$-convergence follows straight-forward by $(a.s.)^*$-convergence. These facts complete the proof.

□

In order to prove essential inequality (4) in the preceding proof and obtain some more details in this direction we shall first recall a general fact saying that *measure preserving transformations are perfect*, see [20]. In particular, given stationary $\xi = \{ \xi_j \mid j \geq 1 \}$ on $(\Omega, \mathcal{F}, P)$ with values in $(S, \mathcal{A})$, the unilateral shift θ is P_ξ-perfect in $(S^{\mathbf{N}}, \mathcal{A}^{\mathbf{N}})$. In other words for any function Ψ from $S^{\mathbf{N}}$ into $\bar{\mathbf{R}}$ we have $(\Psi \circ \theta)^* = \Psi^* \circ \theta$. The next proposition finishes the proof of the preceding theorem.

Proposition 2.

Under the hypothesis in theorem 1 let us suppose that $\| f \|_T$ *belongs to* $L^{<1>}(\pi)$ *and that the map* $s \mapsto f(s,t)$ *is* π*-measurable for every* $t \in T$ *. If* ξ *is* P*-perfect as a map from* Ω *into* $S^{\mathbf{N}}$ *, then for any* A *subset of* T *the following three statements are satisfied:*

$$
\text{(1)} \qquad \limsup_{n\to\infty} \Big(\sup_{t',t''\in A} \mid \frac{1}{n} S_n\big(f(t')-f(t'')\big) \mid \Big)^* = C \quad P\text{-}a.s.
$$

$$
\text{(2)} \qquad C \leq \inf_{n\geq 1} \int^* \sup_{t',t''\in A} \mid \frac{1}{n} S_n\big(f(t')-f(t'')\big) \mid dP
$$

$$
\text{(3)} \qquad \inf_{n\geq 1} \int^* \sup_{t',t''\in A} \mid \frac{1}{n} S_n\big(f(t')-f(t'')\big) \mid dP =
$$
$$
= \limsup_{n\to\infty} \int^* \sup_{t',t''\in A} \mid \frac{1}{n} S_n\big(f(t')-f(t'')\big) \mid dP
$$

where C *is a real constant depending on* A *.*

Proof. It is not a restriction to assume that A equals T, and that the expression $\sup_{t',t''\in A} \mid n^{-1}S_n\big(f(t')-f(t'')\big) \mid$ including differences $f(t')-f(t'')$ in (1), (2) and (3) is replaced by the expression $\sup_{t\in A} \mid n^{-1}S_n(f(t)) \mid$ including a single function $f(t)$. Note that our hypothesis remain valid after this change. For every $n\geq 1$ define a map $\Psi_n : S^{\mathbf{N}} \rightarrow \bar{\mathbf{R}}$ by:

$$\Psi_n(s) \;=\; \sup_{t\in T} \mid \frac{1}{n}\sum_{j=1}^{n} f(s_j,t) \mid \;=\; \| \frac{1}{n}\sum_{j=1}^{n} f(s_j) \|_T$$

for $s=(s_1,s_2\dots)\in S^{\mathbf{N}}$. For given $d\geq 1$, and for $n>d$ put $\sigma_n=[n/d]$ to be the integer part of n/d. Then we have:

$$\Psi_n(s) \;\leq\; \frac{\sigma_n\cdot d}{n}\cdot\frac{1}{\sigma_n}\sum_{j=1}^{\sigma_n}\big(\Psi_d\circ\theta^{(j-1)d}\big)(s) \;+\; \frac{1}{n}\sum_{\sigma_n\cdot d<j\leq n} \| f(s_j) \|_T \tag{4}$$

for all $s=(s_1,s_2\dots)\in S^{\mathbf{N}}$. Indeed let us note that we have:

$$\begin{aligned}\frac{1}{n}\sum_{j=1}^{n} f(s_j) \;&=\; \frac{1}{n}\sum_{j=1}^{\sigma_n\cdot d} f(s_j) \;+\; \frac{1}{n}\sum_{\sigma_n\cdot d<j\leq n} f(s_j)\\ &=\; \frac{1}{n}\sum_{j=1}^{\sigma_n}\sum_{i=1}^{d} f(s_{i+(j-1)d}) \;+\; \frac{1}{n}\sum_{\sigma_n\cdot d<j\leq n} f(s_j)\\ &=\; \frac{\sigma_n\cdot d}{n}\cdot\frac{1}{\sigma_n}\sum_{j=1}^{\sigma_n}\Big(\frac{1}{d}\sum_{i=1}^{d} f(s_{i+(j-1)d})\Big) \;+\; \frac{1}{n}\sum_{\sigma_n\cdot d<j\leq n} f(s_j)\end{aligned}$$

for all $s=(s_1,s_2\dots)\in S^{\mathbf{N}}$. Therefore (4) follows easily by taking supremum over all $t\in T$. And taking the upper P_ξ-envelopes of Ψ_n and Ψ_d in (4) we get:

$$\Psi_n^*(s) \;\leq\; \frac{\sigma_n\cdot d}{n}\cdot\frac{1}{\sigma_n}\sum_{j=1}^{\sigma_n}\big(\Psi_d^*\circ\theta^{(j-1)d}\big)(s) \;+\; \frac{1}{n}\sum_{\sigma_n\cdot d<j\leq n} \| f(s_j) \|_T^*$$

for all $s=(s_1,s_2\dots)\in S^{\mathbf{N}}$, where $\| f \|_T^*$ denotes the upper π-envelope of $\| f \|_T$ as a function from S into $\bar{\mathbf{R}}$. Hence by P-perfectness of ξ we directly find:

$$\begin{aligned}\| \frac{1}{n}\sum_{j=1}^{n} f(\xi_j) \|_T^* \;=\; \Psi_n^*(\xi) \;\leq\; &\frac{\sigma_n\cdot d}{n}\cdot\frac{1}{\sigma_n}\sum_{j=1}^{\sigma_n}\big(\Psi_d^*\circ\theta^{(j-1)d}\big)(\xi)\\ &+\frac{1}{n}\sum_{\sigma_n\cdot d<j\leq n} \| f \|_T^*\circ\xi_j \;.\end{aligned} \tag{5}$$

Now we would like to apply the law of large numbers for stationary sequences of random variables, see (2.5). By (2.3) and (2.4) we may easily deduce that the sequence $\{\, \big(\Psi_d^*\circ\theta^{(j-1)d}\big)(\xi) \mid j\geq 1 \,\}$

is stationary and ergodic. Moreover since $\| f \|_T$ belongs to $L^{<1>}(\pi)$, then we obviously have:

$$\int (\Psi_d^* \circ \xi)\ dP \ \leq \ \int^* \| f(s) \|_T \ \ \pi(ds) \ < \ \infty \ .$$

Thus the law of large numbers for stationary sequences may be applied, and since $(\sigma_n \cdot d)/n \rightarrow 1$ as $n \rightarrow \infty$, then by (2.5) we may conclude:

$$\frac{\sigma_n \cdot d}{n} \cdot \frac{1}{\sigma_n} \sum_{j=1}^{\sigma_n} \big(\Psi_d^* \circ \theta^{(j-1)d}\big)(\xi) \ \longrightarrow \int_\Omega (\Psi_d^* \circ \xi)\ dP \quad P\text{-}a.s. \tag{6}$$

as $n \rightarrow \infty$. And since ξ is by our assumption P-perfect, then we have:

$$\int_\Omega (\Psi_d^* \circ \xi)\ dP \ = \ \int_\Omega^* \| \frac{1}{d} \sum_{j=1}^{d} f(\xi_j) \|_T \ dP \ . \tag{7}$$

Similarly by (2.3) and (2.4) we may conclude that the sequence $\{ \ \| f \|_T^* \circ \xi_j \ \mid \ j \geq 1 \ \}$ is stationary and ergodic. Moreover since $\| f \|_T$ belongs to $L^{<1>}(\pi)$, then we obviously have:

$$\int_\Omega^* \| f \|_T^* \circ \xi_1 \ dP \ = \ \int^* \| f(s) \|_T \ \ \pi(ds) \ < \ \infty \ .$$

Thus the law of large numbers for stationary sequences may be applied, and since $(\sigma_n \cdot d)/n \rightarrow 1$ as $n \rightarrow \infty$, then by (2.5) we may conclude:

$$\begin{aligned} \frac{1}{n} \sum_{\sigma_n \cdot d < j \leq n} \| f \|_T^* \circ \xi_j \ &= \ \frac{1}{n} \sum_{j=1}^{n} \| f \|_T^* \circ \xi_j \ - \ \frac{1}{n} \sum_{j=1}^{\sigma_n \cdot d} \| f \|_T^* \circ \xi_j \\ &= \ \frac{1}{n} \sum_{j=1}^{n} \| f \|_T^* \circ \xi_j \ - \ \frac{\sigma_n \cdot d}{n} \cdot \frac{1}{\sigma_n \cdot d} \sum_{j=1}^{\sigma_n \cdot d} \| f \|_T^* \circ \xi_j \ \rightarrow \ 0 \quad P\text{-}a.s. \end{aligned} \tag{8}$$

as $n \rightarrow \infty$. Now by (5), (6), (7) and (8) we may conclude:

$$\limsup_{n\rightarrow\infty} \ \| \frac{1}{n} \sum_{j=1}^{n} f(\xi_j) \|_T^* \ \leq \ \int_\Omega^* \| \frac{1}{d} \sum_{j=1}^{d} f(\xi_j) \|_T \ dP \quad P\text{-}a.s. \tag{9}$$

for all $d \geq 1$. Therefore by showing (1), statement (2) will be also established. And for (1), since by our assumption ξ is P-perfect, it is enough to show that we have:

$$\limsup_{n\rightarrow\infty} \ \Psi_n^*(\xi) \ = \ C \quad P\text{-}a.s. \tag{10}$$

In order to establish (10) it is enough to show that the map $\limsup_{n\rightarrow\infty} \ \Psi_n^*$ is θ-invariant mod P_ξ , see [17] (p.5), that is:

$$(11) \qquad \limsup_{n\to\infty} \ \Psi_n^* \circ \theta \ = \ \limsup_{n\to\infty} \ \Psi_n^* \quad P_\xi\text{-}a.s.$$

Since θ is P_ξ-perfect, then we have $\Psi_n^* \circ \theta = (\Psi_n \circ \theta)^*$. Moreover we also have:

$$(\Psi_n \circ \theta) \ = \ \sup_{t\in T} \ \Big| \ \frac{1}{n} \sum_{j=1}^{n} f(s_{j+1}, t) \ \Big| \ = \ \sup_{t\in T} \ \Big| \ \frac{n+1}{n} \cdot \frac{1}{n+1} \sum_{j=1}^{n+1} f(s_j, t) \ - \ \frac{1}{n} \cdot f(s_1, t) \ \Big|$$

for all $s = (s_1, s_2 \dots) \in S^{\mathbf{N}}$. Using these two facts one can easily verify the validity of (11). Note since $\| \ f \ \|_T$ belongs to $L^{<1>}(\pi)$ that we have $\| \ f \ \|_T < \infty \ \ \pi\text{-}a.s.$. Thus $\limsup_{n\to\infty} \ \Psi_n^*$ is θ-invariant mod P_ξ , and (10) follows by the ergodicity of ξ . To prove (3) take the P-integral on both sides in (5), then by (7) we have:

$$\int^* \| \ \frac{1}{n} \sum_{j=1}^{n} f(\xi_j) \ \|_T \ dP \ \leq \ \frac{\sigma_n \cdot d}{n} \int_\Omega^* \| \ \frac{1}{d} \sum_{j=1}^{d} f(\xi_j) \ \|_T \ dP \ + \ \frac{n - \sigma_n \cdot d}{n} \int^* \| \ f \ \|_T \ \ \pi(ds) \ .$$

Since $(\sigma_n \cdot d)/n \leq 1$ and $(n - \sigma_n \cdot d)/n \leq d/n$, and since by our assumption $\| \ f \ \|_T$ belongs to $L^{<1>}(\pi)$, we may obtain:

$$\limsup_{n\to\infty} \ \int^* \| \ \frac{1}{n} \sum_{j=1}^{n} f(\xi_j) \ \|_T \ dP \ \leq \ \int^* \| \ \frac{1}{d} \sum_{j=1}^{d} f(\xi_j) \ \|_T \ dP$$

for all $d \geq 1$. Now (3) follows straight forward by taking infimum over all $d \geq 1$, and the proof is complete.

□

Remark 3.

Under the hypothesis in theorem 1 and proposition 2 it is easily verified that in the case where the map $(s_1, s_2 \dots) \mapsto \sup_{t', t'' \in A} \ | \ n^{-1} \sum_{j=1}^{n} (f(s_j, t') - f(s_j, t'')) \ |$ is P_ξ-measurable for all $n \geq 1$ the assumption of perfectness on ξ is not needed for their conclusions remain valid.

Theorem 4.

Let $\xi = \{ \ \xi_j \ | \ j \geq 1 \ \}$ *be a stationary ergodic sequence of random variables defined on a probability space* $(\Omega, \mathcal{F}, P)$ *with values in a measurable space* $(S, \mathcal{A})$ *and a common distribution law* π *, let* T *be an arbitrary set, and let* $f : S \times T \to \mathbf{R}$ *be a given map. Let us suppose that* $\| \ f \ \|_T$ *belongs to* $L^{<1>}(\pi)$ *and that the map* $s \mapsto f(s,t)$ *is* π*-measurable for every* $t \in T$ *, and let* $M(t) = \int_S f(s,t) \ \pi(ds)$ *be the* π*-mean function of* f *for* $t \in T$ *. If we have:*

$$(1) \qquad \| \ \frac{1}{n} \sum_{j=1}^{n} f(\xi_j) - M \ \|_T \to 0 \quad (P^*)$$

as $n \to \infty$ *, then* f *is eventually totally bounded in* π*-mean.*

Proof. Let us for $n \geq 1$ and $\alpha \in \Gamma(T)$ put:

$$D_{n,\alpha} = \Big(\max_{A\in\alpha} \sup_{t',t''\in A} \mid \frac{1}{n} S_n\big(f(t') - f(t'')\big) \mid \Big)^* = \max_{A\in\alpha} \Big(\sup_{t',t''\in A} \mid \frac{1}{n} S_n\big(f(t') - f(t'')\big) \mid \Big)^* .$$

Then we obviously have:

$$D_{n,\alpha} \; \leq \; 2 \parallel \frac{1}{n} S_n(f) \parallel_T^* \; \leq \; \frac{2}{n} \sum_{j=1}^{n} \parallel f \parallel_T^* \circ \xi_j$$

for all $n \geq 1$ and all $\alpha \in \Gamma(T)$. Therefore by (2.3) and (2.6) we may conclude that the family of random variables $\{ \, D_{n,\alpha} \mid n \geq 1 \, , \, \alpha \in \Gamma(T) \, \}$ is uniformly integrable. Thus for given $\epsilon > 0$, there exists $0 < \delta < \epsilon/2$ such that:

$$\int_F D_{n,\alpha} \, dP < \epsilon/2 \tag{2}$$

for all $F \in \mathcal{F}$ satisfying $P(F) < \delta$, whenever $n \geq 1$ and $\alpha \in \Gamma(T)$. Since under our assumptions M belongs to $B(T)$, then for given $\delta > 0$, there exists $\alpha_\delta \in \Gamma(T)$ such that:

$$\sup_{t',t''\in A} \mid M(t') - M(t'') \mid \; \leq \; \delta/2 \tag{3}$$

for all $A \in \gamma_\delta$. Thus by (1) and (3) we may easily establish:

$$\begin{aligned} P\{ \, D_{n,\gamma_\delta} > \delta \, \} \; &\leq \; P^*\{ \, 2 \parallel \frac{1}{n} S_n(f) - M \parallel_T \; + \\ &\quad + \max_{A\in\gamma_\delta} \sup_{t',t''\in A} \mid M(t') - M(t'') \mid > \delta \, \} \\ &\leq \; P^*\{ \, \parallel \frac{1}{n} S_n(f) - M \parallel_T > \delta/4 \, \} < \delta \end{aligned} \tag{4}$$

for all $n \geq n_\delta$ with some $n_\delta \geq 1$. Now by (2) and (4) we may conclude:

$$\int D_{n,\gamma_\delta} \, dP \; = \; \int_{\{D_{n,\gamma_\delta} \leq \delta\}} D_{n,\gamma_\delta} \, dP \; + \; \int_{\{D_{n,\gamma_\delta} > \delta\}} D_{n,\gamma_\delta} \, dP \; \leq \; \delta + \epsilon/2 \; < \epsilon$$

for all $n \geq n_\delta$. Hence we get:

$$\inf_{n\geq 1} \int^* \sup_{t',t''\in A} \mid \frac{1}{n} S_n\big(f(t') - f(t'')\big) \mid \; dP \; \leq \; \limsup_{n\to\infty} \int D_{n,\gamma_\delta} \, dP < \epsilon$$

for all $A \in \gamma_\delta$. That means f is eventually totally bounded in π-mean, and the proof is complete. □

Corollary 5.

Under the hypothesis in theorem 4 let us suppose that $\parallel f \parallel_T$ *belongs to* $L^{<1>}(\pi)$ *and that the map* $s \mapsto f(s,t)$ *is* π*-measurable for every* $t \in T$ *, and let* $M(t) = \int_S f(s,t) \, \pi(ds)$ *be* *the π-mean function of f for $t \in T$. If ξ is P-perfect as a map from Ω into $S^{\mathbf{N}}$, then the following four statements are equivalent:*

(1) $\quad \| \, n^{-1} S_n(f) - M \, \|_T \to 0 \quad (a.s.)^*$

(2) $\quad \| \, n^{-1} S_n(f) - M \, \|_T \to 0 \quad (L^1)^*$

(3) $\quad \| \, n^{-1} S_n(f) - M \, \|_T \to 0 \quad (P^*)$

(4) *The map f is eventually totally bounded in π-mean.*

Proof. Straight forward by theorem 1 and theorem 4 using (2.12).

□

The next proposition shows that $(L^1)^*$- and (P^*)-convergence in the preceding corollary may be apparently relaxed.

Proposition 6.

Under the hypothesis in theorem 4 let us suppose that $\| \, f \, \|_T$ belongs to $L^{<1>}(\pi)$ and that the map $s \mapsto f(s,t)$ is π-measurable for every $t \in T$, and let $M(t) = \int_S f(s,t) \, \pi(ds)$ be the π-mean function of f for $t \in T$. If ξ is P-perfect as a map from Ω into $S^{\mathbf{N}}$, then $\| \, n^{-1} S_n(f) - M \, \|_T \to 0 \;\; (P^*)$, if and only if we have:

$$\inf_{n \ge 1} \; P^*\{ \, \| \, \frac{1}{n} S_n(f) - M \, \|_T > \epsilon \, \} < \epsilon \tag{1}$$

for every $\epsilon > 0$. Similarly, if ξ is P-perfect as a map from Ω into $S^{\mathbf{N}}$, then $\| \, n^{-1} S_n(f) - M \, \|_T \to 0 \;\; (L^1)^*$, if and only if we have:

$$\inf_{n \ge 1} \; \int^* \| \, \frac{1}{n} S_n(f) - M \, \|_T \; dP \; = \; 0 \; . \tag{2}$$

Proof. First we show that (1) implies (2). So suppose that (1) holds, and let us put:

$$D_n \; = \| \, \frac{1}{n} S_n(f) - M \, \|_T^*$$

for all $n \ge 1$. Then we have:

$$D_n \; \le \; \frac{1}{n} \sum_{j=1}^{n} \| \, f \, \|_T^* \circ \xi_j \; + \; \| \, M \, \|_T$$

for all $n \ge 1$. Since by our hypothesis $\| \, f \, \|_T$ belongs to $L^{<1>}(\pi)$ and M belongs to $B(T)$, then by (2.3) and (2.6) the sequence of random variables $\{ \, D_n \mid n \ge 1 \, \}$ is uniformly integrable. Thus for given $\epsilon > 0$, there exists $0 < \delta < \epsilon/2$ such that:

$$\int_F D_n \; dP < \epsilon/2 \tag{3}$$

whenever $F \in \mathcal{F}$ and $P(F) < \delta$, for all $n \ge 1$. According to (1) for given $\delta > 0$, there exists $n_\delta \geq 1$ such that:

$$P\{\ D_{n_\delta} > \delta\ \} < \delta \tag{4}$$

Now by (3) and (4) we may conclude:

$$\int^* \| \frac{1}{n_\delta} S_{n_\delta}(f) - M \|_T \ dP \ = \ \int D_{n_\delta} \ dP \ = \ \int_{\{D_{n_\delta} \leq \delta\}} D_{n_\delta} \ dP$$

$$+ \int_{\{D_{n_\delta} > \delta\}} D_{n_\delta} \ dP \ \leq \ \delta \ + \ \epsilon/2 \ < \ \epsilon \ .$$

This fact establishes (2). Next we show that (2) implies given $(L^1)^*$-convergence. For this one can easily verify that the same arguments as those given in the proof of proposition 2 apply also here, and in this way we could obtain:

$$\inf_{n \geq 1} \int^* \| \frac{1}{n} S_n(f) - M \|_T \ dP \ = \ \limsup_{n \to \infty} \int^* \| \frac{1}{n} S_n(f) - M \|_T \ dP \ .$$

Therefore (2) provides $\| n^{-1} S_n(f) - M \|_T^* \to 0$ for $n \to \infty$. Since in general $(L^1)^*$-convergence implies (P^*)-convergence, see (2.12), these facts complete the proof.

$\square$

We pass to a weak convergence. Let us consider a Banach space B and a sequence of arbitrary functions $\{ \zeta_j \mid j \geq 1 \}$ from a probability space $(\Omega, \mathcal{F}, P)$ into B . Let $C(B)$ denote the set of all bounded continuous functions from B into $\mathbf{R}$, and let $\mathcal{K}(B)$ denote the family of all compact subsets of B . Let τ be a probability measure defined on the Borel σ-algebra in B , then the sequence $\{ \zeta_j \mid j \geq 1 \}$ is said to be:

(i) *weakly convergent* to τ , if we have:

$$\lim_{n \to \infty} \int^* F(\zeta_n) \ dP \ = \ \lim_{n \to \infty} \int_* F(\zeta_n) \ dP \ = \ \int_B F \ d\tau$$

for all $F \in C(B)$, and in this case we shall write $\zeta_n \to \tau$ *weakly in* B

(ii) *uniformly tight*, if $\forall \epsilon > 0$, $\exists K_\epsilon \in \mathcal{K}(B)$ such that:

$$\limsup_{n \to \infty} P^*\{\ \zeta_n \notin K_\epsilon\ \} \leq \epsilon$$

(iii) *eventually tight*, if $\forall \epsilon > 0$, $\exists K_\epsilon \in \mathcal{K}(B)$ such that:

$$\limsup_{n \to \infty} \int^* F(\zeta_n) \ dP \leq \epsilon$$

for all $F \in C(B)$ satisfying $0 \leq F \leq 1_{B \setminus K_\epsilon}$.

It is easy to verify that we have, see [15]:

(8) If $\{ \zeta_j \mid j \geq 1 \}$ is uniformly tight, it is also eventually tight.

(9) If $\zeta_n \to \tau$ weakly in B and τ is Radon, then $\{ \zeta_j \mid j \geq 1 \}$ is eventually tight.

(10) If $\zeta_n \to c \ (P^*)$ for some $c \in B$, then $\zeta_n \to c$ weakly in B .

Clarify that $\zeta_n \to c \ (P^*)$ means $\| \zeta_n - c \| \to 0 \ (P^*)$ where $\| \cdot \|$ denotes the norm in B , and that $\zeta_n \to c$ weakly in B means $\zeta_n \to \delta_c$ where δ_c denotes the Dirac measure concentrated at the point c . And our next aim is to show that the eventual tightness of the sequence of averages $\{ n^{-1} S_n(f) \mid n \geq 1 \}$ in the Banach space $(B(T), \| \cdot \|_T)$ is equivalent in the stationary case to the uniform law of large numbers. One implication in this direction is obvious. Namely if $\| n^{-1} S_n(f) - M \|_T \to 0 \ (P^*)$ as $n \to \infty$, then by (9) and (10) above we may conclude that the sequence $\{ n^{-1} S_n(f) \mid n \geq 1 \}$ is eventually tight. The rest is contained in the following equivalently formulated theorem.

Theorem 7.

Let $\xi = \{ \xi_j \mid j \geq 1 \}$ *be a stationary ergodic sequence of random variables defined on a probability space* $(\Omega, \mathcal{F}, P)$ *with values in a measurable space* $(S, \mathcal{A})$ *and a common distribution law* π *, let* T *be an arbitrary set, and let* $f : S \times T \to \mathbf{R}$ *be a given map. Let us suppose that* $\| f \|_T$ *belongs to* $L^{<1>}(\pi)$ *and that the map* $s \mapsto f(s,t)$ *is* π*-measurable for every* $t \in T$ *. If* ξ *is* P*-perfect as a map from* Ω *into* $S^{\mathbf{N}}$ *, then the following two statements are equivalent:*

(1) *The map* f *is eventually totally bounded in* π*-mean*

(2) *The sequence of averages* $A(f) = \{ n^{-1} S_n(f) \mid n \geq 1 \}$ *is eventually tight in* $B(T)$ *.*

Proof. The implication $(1) \Rightarrow (2)$ follows by theorem 1, (2.12) and (9)+(10) above as we already mentioned. For reverse implication suppose that (2) holds. Then we claim that for given $\delta > 0$, there exists $\gamma_\delta \in \Gamma(T)$ such that:

$$P^* \{ \max_{A \in \gamma_\delta} \sup_{t',t'' \in A} \mid \frac{1}{n} S_n\big(f(t') - f(t'')\big) \mid > \delta \} < \delta \tag{3}$$

for all $n \geq n_\delta$ with some $n_\delta \geq 1$. To prove (3) we may proceed as follows. Since $\{ n^{-1} S_n(f) \mid n \geq 1 \}$ is eventually tight in $B(T)$, then for given $\delta > 0$, there exists $K_\delta \in \mathcal{K}(B(T))$ such that:

$$\limsup_{n \to \infty} \int^* F\big(n^{-1} S_n(f)\big) \, dP \ \leq \ \delta \tag{4}$$

for all $F \in C(B(T))$ satisfying $0 \leq F \leq 1_{B(T) \setminus K_\delta}$. The compactness of K_δ yields the existence of $\gamma_\delta \in \Gamma(T)$ satisfying:

$$\sup_{t',t'' \in A} \mid \varphi(t') - \varphi(t'') \mid < \delta/3 \tag{5}$$

for all $\varphi \in K_\delta$ and all $A \in \gamma_\delta$, see [7] (p.260). By using (5) one can easily verify that for any $\psi \in b(K_\delta, \delta/3) = \bigcup_{\varphi \in K_\delta} b(\varphi, \delta/3)$ we have:

$$\sup_{t',t'' \in A} \mid \psi(t') - \psi(t'') \mid < \delta \tag{6}$$

for all $A \in \gamma_\delta$. Now by (4) and (6) we may conclude:

$$\begin{aligned} P^*\{ \max_{A \in \gamma_\delta} \sup_{t',t'' \in A} &\mid \frac{1}{n} S_n\big(f(t') - f(t'')\big) \mid > \delta \} \\ &= P^*\{ \max_{A \in \gamma_\delta} \sup_{t',t'' \in A} \mid \frac{1}{n} S_n\big(f(t')\big) - \frac{1}{n} S_n\big(f(t'')\big) \mid > \delta \} \\ &\le \int^* 1_{\{n^{-1}S_n(f) \notin b(K_\delta, \delta/3)\}} \, dP \;\le\; \int^* F_\delta\big(n^{-1}S_n(f)\big) \, dP \;\le\; \delta \end{aligned}$$

for all $n \ge n_\delta$ with some $n_\delta \ge 1$, where $F_\delta \in C(B(T))$ is chosen to satisfy $F_\delta(\varphi) = 1$ for $\varphi \in B(T) \setminus b(K_\delta, \delta/3)$, $F_\delta(\varphi) = 0$ for $\varphi \in K_\delta$, and $0 \le F_\delta(\varphi) \le 1_{B(T)\setminus K_\delta}(\varphi)$ for all $\varphi \in B(T)$. These facts complete the proof of (3). Let us in addition for $n \ge 1$ and $\alpha \in \Gamma(T)$ put:

$$\begin{aligned} D_{n,\alpha} &= \Big(\max_{A \in \alpha} \sup_{t',t'' \in A} \mid \frac{1}{n} S_n\big(f(t') - f(t'')\big) \mid \Big)^* \\ &= \max_{A \in \alpha} \Big(\sup_{t',t'' \in A} \mid \frac{1}{n} S_n\big(f(t') - f(t'')\big) \mid \Big)^* . \end{aligned}$$

Then by (2.3) and (2.6) we may conclude as in the proof of theorem 4 that the family $\{ D_{n,\alpha} \mid n \ge 1 , \alpha \in \Gamma(T) \}$ is uniformly integrable. Therefore for given $\epsilon > 0$, there exists $0 < \delta < \epsilon/2$ such that:

$$\int_F D_{n,\alpha} \, dP < \epsilon/2 \tag{7}$$

whenever $F \in \mathcal{F}$ and $P(F) < \delta$, for all $n \ge 1$ and all $\alpha \in \Gamma(T)$. Let us now for given $\delta > 0$ apply (3) and find $\gamma_\delta \in \Gamma(T)$ satisfying:

$$P\{ D_{n,\gamma_\delta} > \delta \} < \delta . \tag{8}$$

Now by (7) and (8) we may obtain:

$$\int D_{n,\gamma_\delta} \, dP = \int_{\{D_{n,\gamma_\delta} \le \delta\}} D_{n,\gamma_\delta} \, dP + \int_{\{D_{n,\gamma_\delta} > \delta\}} D_{n,\gamma_\delta} \, dP \le \delta + \epsilon/2 < \epsilon$$

for all $n \ge 1$. Therefore we may conclude:

$$\inf_{n \ge 1} \int^* \sup_{t',t'' \in A} \mid \frac{1}{n} S_n\big(f(t') - f(t'')\big) \mid \, dP \le \limsup_{n \to \infty} \int D_{n,\gamma_\delta} \, dP < \epsilon \tag{9}$$

for all $A \in \gamma_\delta$, and the proof is complete.

□

Corollary 8.

Under the hypothesis in corollary 5 let us suppose that $\| f \|_T$ *belongs to* $L^{<1>}(\pi)$ *and that the map* $s \mapsto f(s,t)$ *is* π*-measurable for every* $t \in T$ *, and let* $M(t) = \int_S f(s,t)\, \pi(ds)$ *be the* π*-mean function of* f *for* $t \in T$ *. If* ξ *is* P*-perfect as a map from* Ω *into* $S^{\mathbf{N}}$ *, then the statements (1)-(4) in corollary 5 are also equivalent to the following statement:*

(1) $$n^{-1} S_n(f) \to M \quad \textit{weakly in} \quad B(T)$$

as $n \to \infty$.

Proof. Straight forward by corollary 5, theorem 7 and (9)+(10).

$\square$

We proceed by considering eventually total boundedness in mean in some more technical details.

Theorem 9.

Under the hypothesis in theorem 7 let us suppose that $\| f \|_T$ *belongs to* $L^{<1>}(\pi)$ *and that the map* $s \mapsto f(s,t)$ *is* π*-measurable for every* $t \in T$ *. If* ξ *is* P*-perfect as a map from* Ω *into* $S^{\mathbf{N}}$ *, then the following two statements are equivalent:*

(1) *The map* f *is eventually totally bounded in* π*-mean*

(2) *For every* $\epsilon > 0$ *, there exists* $\gamma_\epsilon \in \Gamma(T)$ *such that either of the following seven conditions is satisfied:*

(i) $$\inf_{n \geq 1} \int^* \max_{A \in \gamma_\epsilon} \sup_{t',t'' \in A} \Big| \frac{1}{n} S_n\big(f(t') - f(t'')\big) \Big| \; dP \; < \epsilon$$

(ii) $$\limsup_{n \to \infty} \int^* \sup_{t',t'' \in A} \Big| \frac{1}{n} S_n\big(f(t') - f(t'')\big) \Big| \; dP \; < \epsilon \;\; , \;\; \forall A \in \gamma_\epsilon$$

(iii) $$\limsup_{n \to \infty} \int^* \max_{A \in \gamma_\epsilon} \sup_{t',t'' \in A} \Big| \frac{1}{n} S_n\big(f(t') - f(t'')\big) \Big| \; dP \; < \epsilon$$

(iv) $$\inf_{n \geq 1} P^* \Big\{ \sup_{t',t'' \in A} \Big| \frac{1}{n} S_n\big(f(t') - f(t'')\big) \Big| > \epsilon \Big\} \; < \; \epsilon \;\; , \;\; \forall A \in \gamma_\epsilon$$

(v) $$\inf_{n \geq 1} P^* \Big\{ \max_{A \in \gamma_\epsilon} \sup_{t',t'' \in A} \Big| \frac{1}{n} S_n\big(f(t') - f(t'')\big) \Big| > \epsilon \Big\} \; < \; \epsilon$$

(vi) $$\limsup_{n \to \infty} P^* \Big\{ \sup_{t',t'' \in A} \Big| \frac{1}{n} S_n\big(f(t') - f(t'')\big) \Big| > \epsilon \Big\} \; < \; \epsilon \;\; , \;\; \forall A \in \gamma_\epsilon$$

(vii) $$\limsup_{n \to \infty} P^* \Big\{ \max_{A \in \gamma_\epsilon} \sup_{t',t'' \in A} \Big| \frac{1}{n} S_n\big(f(t') - f(t'')\big) \Big| > \epsilon \Big\} \; < \; \epsilon.$$

Proof. The implication $(1) \Rightarrow$ (vii) is established in the proof of theorem 7, see its relation (3), as well as the implication $(1) \Rightarrow$ (iii) , see its relation (9). And in order to complete the proof one can easily verify by combining deduced and obvious implications that it is enough to show the implication (iv) $\Rightarrow (1)$. Note that Markov's inequality for upper integrals might be used during this verification, see [20]. So suppose that (iv) holds. Let us for $n \geq 1$, $\alpha \in \Gamma(T)$ and $A \in \Gamma(T)$ put:

$$D_{n,A} \; = \; \Big(\sup_{t',t'' \in A} \mid \frac{1}{n} S_n\big(f(t') - f(t'')\big) \mid \Big)^* \; .$$

Since we have:

$$D_{n,A} \; \leq \; 2 \parallel \frac{1}{n} S_n(f) \parallel_T^* \; \leq \; \frac{2}{n} \sum_{j=1}^{n} \parallel f \parallel_T^* \circ \xi_j$$

for all $n \geq 1$ and all $A \in \alpha$ with $\alpha \in \Gamma(T)$, then by (2.3) and (2.6) we may conclude that the family of random variables $\{ \; D_{n,A} \; \mid \; n \geq 1 \; , \; A \in \alpha \; , \; \alpha \in \Gamma(T) \; \}$ is uniformly integrable. Therefore for given $\epsilon > 0$, there exists $0 < \delta < \epsilon/2$ such that:

$$\int_F D_{n,A} \; dP \; < \epsilon/2 \tag{3}$$

whenever $F \in \mathcal{F}$ and $P(F) < \delta$, for all $n \geq 1$ and all $A \in \alpha$ with $\alpha \in \Gamma(T)$. And by (iv) for given $\delta > 0$, there exist $\gamma_\delta \in \Gamma(T)$ and $n_\delta \geq 1$ such that:

$$P\{ \; D_{n_\delta,A} > \delta \; \} < \delta \tag{4}$$

for all $A \in \gamma_\delta$. Now by (3) and (4) we may conclude:

$$\int^* \sup_{t',t'' \in A} \mid \frac{1}{n_\delta} S_{n_\delta}\big(f(t') - f(t'')\big) \mid \; dP \; = \; \int D_{n_\delta,A} \; dP$$

$$= \; \int_{\{D_{n_\delta,A} \leq \delta\}} D_{n_\delta,A} \; dP \; + \; \int_{\{D_{n_\delta,A} > \delta\}} D_{n_\delta,A} \; dP \; \leq \; \delta + \epsilon/2 \; < \epsilon$$

for all $A \in \gamma_\delta$. This fact establishes (1), and the proof is complete.

$\square$

We conclude this paper with a characterization of eventually total boundedness in mean in terms of the existence of certain totally bounded pseudo-metrics, see [12]. Recall that a pseudo-metric ρ on a set T is said to be *totally bounded*, if T may be covered by finitely many ρ-balls of any given radius $r > 0$. And a pseudo-metric ρ on a set T is called *an ultra-pseudo-metric*, if $\rho(s,t) \leq \rho(s,u) \vee \rho(u,t)$ for all $s,t,u \in T$.

Theorem 10.

Let $\xi = \{ \; \xi_j \; \mid \; j \geq 1 \; \}$ *be a stationary ergodic sequence of random variables defined on a probability space* $(\Omega, \mathcal{F}, P)$ *with values in a measurable space* $(S, \mathcal{A})$ *and a common distribution law* π *, let* T *be an arbitrary set, and let* $f : S \times T \to \mathbf{R}$ *be a given map. Let us suppose that* $\parallel f \parallel_T$ *belongs to* $L^{<1>}(\pi)$ *and that the map* $s \mapsto f(s,t)$ *is* π*-measurable for every* $t \in T$ *,* *and let* $M(t) = \int_S f(s,t)\, \pi(ds)$ *be the* π*-mean function of* f *for* $t \in T$ *. If* ξ *is* P*-perfect as a map from* Ω *into* $S^{\mathbf{N}}$ *, then the following three statements are equivalent:*

(1) *The map* f *is eventually totally bounded in* π*-mean*

(2) *There exists a totally bounded ultra-pseudo-metric* ρ *on* T *such that either of the following two equivalent conditions is satisfied:*

(i) $$\lim_{r \downarrow 0} \limsup_{n \to \infty} \int^* \sup_{\rho(t',t'')<r} \Big| \frac{1}{n} S_n\big(f(t') - f(t'')\big) \Big| \; dP \;=\; 0$$

(ii) $$\lim_{r \downarrow 0} \inf_{n \geq 1} \int^* \sup_{\rho(t',t'')<r} \Big| \frac{1}{n} S_n\big(f(t') - f(t'')\big) \Big| \; dP \;=\; 0$$

(3) *For every* $\epsilon > 0$ *there exist a totally bounded pseudo-metric* ρ_ϵ *on* T *and* $r_\epsilon > 0$ *such that either of the following two equivalent conditions is satisfied:*

(i) $$\limsup_{n \to \infty} P^*\{ \sup_{\rho_\epsilon(s,t)<r_\epsilon} \Big| \frac{1}{n} S_n\big(f(s) - f(t)\big) \Big| > \epsilon \} < \epsilon \;\;,\quad \forall t \in T$$

(ii) $$\inf_{n \geq 1} P^*\{ \sup_{\rho_\epsilon(s,t)<r_\epsilon} \Big| \frac{1}{n} S_n\big(f(s) - f(t)\big) \Big| > \epsilon \} < \epsilon \;\;,\quad \forall t \in T \;.$$

Proof. We first prove that (1) implies (i) in (2). So suppose that (1) holds. Let $\{ r_n \mid n \geq 0 \}$ be a sequence of real numbers satisfying $1 = r_0 > r_1 > r_2 > \ldots > 0$ with $\lim_{n \to \infty} r_n = 0$. Then for every $n \geq 1$, there exists $\gamma_n \in \Gamma(T)$ such that $| M(t') - M(t'') | < (r_{n-1})^2$ for all $t', t'' \in A$ and all $A \in \gamma_n$. It is not a restriction to assume that $\gamma_1 \subset \gamma_2 \subset \ldots$. Let us define:

$$\rho(t',t'') \;=\; \sup_{n \geq 1} \Big(r_{n-1} \cdot \max_{A \in \gamma_n} | 1_A(t') - 1_A(t'') | \;\Big)$$

for all $t', t'' \in T$. Then ρ is evidently a totally bounded ultra-pseudo-metric on T , and it is easy to verify that $\rho(t',t'') < \epsilon$ implies $| M(t') - M(t'') | < \epsilon^2$ for $0 < \epsilon < 1$ and $t', t'' \in T$. Therefore by Markov's inequality for upper integrals we may conclude:

$$P^*\{ \sup_{\rho(t',t'')<\epsilon} \Big| \frac{1}{n} S_n\big(f(t') - f(t'')\big) \Big| > \epsilon \} \;\leq\; \frac{1}{\epsilon} \int^* \sup_{\rho(t',t'')<\epsilon} \Big| \frac{1}{n} S_n\big(f(t') - f(t'')\big) \Big| \; dP$$

$$\leq\; \frac{1}{\epsilon} \int^* \sup_{\rho(t',t'')<\epsilon} \Big(\; \Big| \frac{1}{n} S_n\big(f(t')\big) - M(t') \Big| \;+\; | M(t') - M(t'') |$$

$$+\; \Big| M(t'') - \frac{1}{n} S_n\big(f(t'')\big) \Big| \; \Big) \; dP \;\leq\; \frac{1}{\epsilon} \Big(\, 2 \int^* \Big\| \, \frac{1}{n} S_n(f) - M \, \Big\|_T \; dP \;+\; \epsilon^2 \Big)$$

for all $0 < \epsilon < 1$ and all $n \geq 1$. Hence by (1) and theorem 1 we can obtain:

(4) $$\limsup_{n \to \infty} P^*\{ \sup_{\rho(t',t'')<\epsilon} \Big| \frac{1}{n} S_n\big(f(t') - f(t'')\big) \Big| > \epsilon \} \leq \epsilon$$

for all $0 < \epsilon < 1$. Let us put:

$$D_{n,\epsilon} \;=\; \Big(\sup_{\rho(t',t'')<\epsilon} \mid \frac{1}{n} S_n\big(f(t') - f(t'')\big) \mid \Big)^*$$

for all $n \geq 1$ and all $\epsilon > 0$. Then we have:

$$D_{n,\epsilon} \;\leq\; \frac{2}{n} \sum_{j=1}^{n} \parallel f \parallel_T^* \circ \xi_j$$

for all $n \geq 1$ and all $\epsilon > 0$. Therefore by (2.3) and (2.6) the family of random variables $\{ D_{n,\epsilon} \mid n \geq 1 , \epsilon > 0 \}$ is uniformly integrable. Thus for given $\epsilon > 0$, there exists $0 < \delta < (\epsilon/2) \wedge 1$ such that:

$$\int_F D_{n,\delta} \; dP \;\leq\; \epsilon/2 \tag{5}$$

whenever $F \in \mathcal{F}$ and $P(F) < \delta$. Now by (4) and (5) we may conclude:

$$\int D_{n,\delta} \; dP \;=\; \int_{\{D_{n,\delta} \leq \delta\}} D_{n,\delta} \; dP \;+\; \int_{\{D_{n,\delta} > \delta\}} D_{n,\delta} \; dP \;\leq\; \delta + \epsilon/2 \;< \epsilon$$

for all $n \geq n_\delta$ with some $n_\delta \geq 1$. Therefore we may obtain:

$$\limsup_{n\to\infty} \int^* \sup_{\rho(t',t'')<\delta} \mid \frac{1}{n} S_n\big(f(t') - f(t'')\big) \mid \; dP \;=\; \limsup_{n\to\infty} \int D_{n,\delta} \; dP \;<\; \epsilon$$

and (i) in (2) is proved. Using the argumentation established in the proof of proposition 2 we may easily obtain:

$$\limsup_{n\to\infty} \int^* \sup_{\rho(t',t'')<r} \mid \frac{1}{n} S_n\big(f(t') - f(t'')\big) \mid \; dP$$

$$= \inf_{n\geq 1} \int^* \sup_{\rho(t',t'')<r} \mid \frac{1}{n} S_n\big(f(t') - f(t'')\big) \mid \; dP$$

for all $r > 0$. Therefore the equivalence between (i) and (ii) in (2) becomes obvious. The implications (i) in $(2) \Rightarrow$ (i) in (3) and (ii) in $(2) \Rightarrow$ (ii) in (3) follow straight forward by applying Markov's inequality for upper integrals. Since (i) in (3) obviously implies (ii) in (3), then it is enough to show that (ii) in (3) implies (1). So suppose that (ii) in (3) holds. Then for given $\epsilon > 0$ there exists a totally bounded pseudo-metric ρ_ϵ on T and $r_\epsilon > 0$ satisfying (ii) in (3). Since ρ_ϵ is totally bounded we can cover the whole set T by finitely many ρ-balls $B_1, B_2, \ldots, B_n$ of given radius r_ϵ . Therefore putting $\gamma_\epsilon = \{ B_1, B_2, \ldots, B_n \}$ we see that (iv) in (2) in theorem 9 is satisfied (with 2ϵ instead of ϵ), and in this way we can obtain (1). These facts complete the proof. □

Acknowledgment. The authors thank Jørgen Hoffmann-Jørgensen and Svend Erik Graversen for valuable discussions on related problems and the great hospitality at University of Aarhus in the spring of 1992.

REFERENCES

[1] AARONSON, J. (1977). On the ergodic theory of non-integrable functions and infinite measure spaces. *Israel J. Math.* 27 (163-173).

[2] BERGER, E. (1990). An almost sure invariance principle for stationary ergodic sequences of Banach space valued random variables. *Probab. Theory Relat. Fields* 84 (161-201).

[3] BREIMAN, L. (1968). *Probability*. Addison-Wesley Publishing Company, Inc.

[4] CHOQUET, G. (1969). *Lectures on Analysis, Vol II: Representation Theory*. W. A. Benjamin, Inc.

[5] DONSKER, M. D. *and* VARADHAN, S. R. S. (1985). Large deviations for stationary Gaussian processes. *Commun. Math. Phys.* 97 (187-210).

[6] DOOB, J. (1953). *Stochastic Processes*. John Wiley and Sons, Inc.

[7] DUNFORD, N. *and* SCHWARTZ, J. T. (1958). *Linear Operators, Part I: General Theory*. Interscience Publ. Inc., New York.

[8] FURSTENBERG, H. (1981). *Recurrence in Ergodic Theory and Combinatorial Number Theory*. Princeton University Press.

[9] GARCIA, A. M. (1970). *Topics in Almost Everywhere Convergence*. Lectures in Advanced Mathematics 4, Markham Publishing Company.

[10] GINÉ, J. and ZINN, J. (1984). Some limit theorems for empirical processes. *Ann. Probab.* 12 (929-989).

[11] HALMOS, P. R. (1956). *Lectures on Ergodic Theory*. The Mathematical Society of Japan.

[12] HOFFMANN-JØRGENSEN, J. (1984). Necessary and sufficient conditions for the uniform law of large numbers. *Proc. Probab. Banach Spaces* V (Medford 1984), *Lecture Notes in Math.* 1153, Springer-Verlag Berlin Heidelberg (258-272).

[13] HOFFMANN-JØRGENSEN, J. (1985). The law of large numbers for non-measurable and non-separable random elements. *Astérisque* 131 (299–356).

[14] HOFFMANN-JØRGENSEN, J. (1990). Uniform convergence of martingales. *Proc. Probab. Banach Spaces* VII (Oberwolfach 1988), *Progr. Probab.* Vol. 21, Birkhäuser Boston, (127-137).

[15] HOFFMANN-JØRGENSEN, J. (1991). *Stochastic Processes on Polish Spaces*. *Various Publ. Ser.* Vol. 39.

[16] HOFFMANN-JØRGENSEN, J. (1992). *Asymptotic Likelihood Theory*. *Various Publ. Ser.* Vol. 40.

[17] KRENGEL, U. (1985). *Ergodic Theorems*. Walter de Gruyter & Co., Berlin.

[18] LEVENTAL, S. (1989). A uniform CLT for uniformly bounded families of martingale

differences. *J. Theor. Probab.* 2, No. 3 (271-287).

[19] MALLER, R. A. (1980). On the law of large numbers for stationary sequences. *Stochastic Processes Appl.* 10 (65-73).

[20] PESKIR, G. (1991). Perfect measures and maps. *Math. Inst. Aarhus, Preprint Series* No. 26 (34 pp).

[21] PESKIR, G. (1991). Vitali convergence theorem for upper integrals. *Math. Inst. Aarhus, Preprint Ser.* No. 28, (14 pp). *Proc. Funct. Anal.* IV (Dubrovnik 1993), *Various Publ. Ser.* Vol. 43, 1994 (191-203).

[22] PESKIR, G. (1992). Consistency of statistical models described by families of reversed submartingales. *Math. Inst. Aarhus, Preprint Ser.* No. 9, (26 pp). *Probab. Math. Statist.* 18, 1998 (289-318).

[23] PESKIR, G. (1992). Uniform convergence of reversed martingales. *Math. Inst. Aarhus, Preprint Ser.* No. 21, (27 pp). *J. Theoret. Probab.* 8, 1995 (387-415).

[24] PETERSEN, K. (1983). *Ergodic Theory*. Cambridge University Press.

[25] ROSALSKY, A. (1987). A strong law for a set-indexed partial sum process with applications to exchangeable and stationary sequences. *Stochastic Processes Appl.* 26 (277-287).

[26] ROZANOV, Yu. A. (1967). *Stationary Random Processes*. Golden-Day, Inc. 500 Sansome Street, San Francisco, California.

[27] SHIRYAYEV, A. N. (1984). *Probability*. Springer-Verlag New York Inc.

[28] STUTE, W. *and* SCHUMANN, G. (1980). A general Glivenko-Cantelli theorem for stationary sequences of random observations. *Scand. J. Statist.* 7 (102-104).

[29] TALAGRAND, M. (1987). The Glivenko-Cantelli problem. *Ann. Probab.* 15 (837-870).

[30] TANNY, D. (1974). A zero-one law for stationary sequences. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* 30 (139-148).

[31] TANNY, D. (1989). A new proof of the strong law without first moment. *Measure and measurable dynamics, Proc. Conf. Rochester*/NY 1987, *Contemp. Math.* 94 (315-320).

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

[33] VAPNIK, V. N. *and* CHERVONENKIS, A. Ya. (1981). Necessary and sufficient conditions for the uniform convergence of means to their expectations. *Theory Probab. Appl.* 26 (532-553).

[34] VOLNY, D. (1987). Martingale decomposition of stationary processes. *Yokohama Math. J.* 35, No. 1/2 (113-121).

[35] WEBER, M. (1988). Stochastic processes with value in exponential type Orlicz spaces. *Ann. Probab.* 16 (1365-1371).

[36] WEBER, M. (1990). Une version fonctionnelle du théorème ergodique poncuel. *C. R. Acad. Sci. Paris, Sér I Math.* 311 (131-133).

Goran Peskir
Department of Mathematical Sciences
University of Aarhus, Denmark
Ny Munkegade, DK-8000 Aarhus
home.imf.au.dk/goran
goran@imf.au.dk

Michel Weber
I.R.M.A. Unité de Recherche
associée C.N.R.S., 1
7, rue René Descartes
67084 Strasbourg
France