The Annals of Mathematical Statistics
1970, Vol. 41, No. 4, 1339–1343

NOTE ON A RESULT OF DUDLEY ON THE SPEED OF MEAN GLIVENKO–CANTELLI CONVERGENCE

BY PETER GÄNSSLER

Universität zu Köln

1. Summary. In a recent paper R. M. Dudley [2] obtained very interesting results on the speed of mean Glivenko–Cantelli convergence where the underlying random variables are supposed to be independent identically distributed (i.i.d.) taking their values in a separable metric space. In the present paper we shall show that his method of proof also applies for the case of Markov processes with stationary transition probabilities fulfilling Doeblin's condition (D_0). The main additional tool in the treatment of this more general case is the determination of an appropriate upper estimate for the variance of the empirical p-measure for the transitions performed in a given set (Lemma 3.2 and Corollary 3.3 (i)) using a mixing property (Lemma 3.1).

As indicated by Dudley, the results obtained in this way are applicable to problems in testing statistical hypotheses. An application concerning the speed of convergence of asymptotically normal estimates for Markov processes will be given in a separate paper [3].

2. Definitions and preliminaries. Let $\mathscr{X}$ be a complete separable metric space, $\mathscr{B}$ the Borel subsets of $\mathscr{X}$ and $(X_n)_{n \in \mathbf{N}}$ a Markov process on some probability space $(\Omega, \mathscr{F}, P)$ with values in $\mathscr{X}$ and stationary transition probabilities $p(\zeta, A) = P\{X_{n+1} \in A \mid X_n = \zeta\}$, $A \in \mathscr{B}$, $\zeta \in \mathscr{X}$, fulfilling Hypothesis (D_0) (see [1] page 221). $\mathbf{N}$ denotes the set of nonnegative integers, and "$:=$" is used for equality by definition. Furthermore, if P is some p-measure (=probability measure) on a σ-field $\mathscr{F}$, we write $P \mid \mathscr{F}$ for short.

According to Kolmogorov's Consistency Theorem ([4], V. 5) we may assume without loss of generality that the measurable space $(\Omega, \mathscr{F})$ is the countable product space $(\mathscr{X}^{\mathbf{N}}, \times_{n \in \mathbf{N}} \mathscr{B})$, $P \mid \times_{n \in \mathbf{N}} \mathscr{B}$ is the p-measure uniquely determined by a given set of transition probabilities $p \mid \mathscr{X} \times \mathscr{B}$ and initial distribution $p_0 \mid \mathscr{B}$, and where $(X_n)_{n \in \mathbf{N}}$ is taken to be the coordinate process.

Let us denote the elements of $\mathscr{X}^{\mathbf{N}}$ by $\mathbf{x}$ and let $\pi_2^{(i)}$, $i \in \mathbf{N}$, be the projection map from $\mathscr{X}^{\mathbf{N}}$ onto $\mathscr{X} \times \mathscr{X}$ defined by $\pi_2^{(i)}(\mathbf{x}) := (x_i, x_{i+1})$. Under Hypothesis (D_0) there exists a unique stationary distribution $p^s \mid \mathscr{B}$ (see [1], V. 5), the corresponding p-measure on $\times_{n \in \mathbf{N}} \mathscr{B}$ determined by $p \mid \mathscr{X} \times \mathscr{B}$ and $p^s \mid \mathscr{B}$ will be denoted by P^s, the two-dimensional marginal distribution pertaining to P^s by $P_2{}^s$.

If $\mathbf{x} = (x_0, x_1, \cdots)$ is an observation on the process, an empirical p-measure (of the sample size n) for the transitions performed in a set $T \in \mathscr{B}_S$, where $\mathscr{B}_S$ denotes the Borel sets of $S := \mathscr{X} \times \mathscr{X}$, is defined by

(2.1) $$Q_n^{\mathbf{x}}(T) := n^{-1} \sum_{i=0}^{n-1} (\chi_T \circ \pi_2^{(i)})(\mathbf{x}),$$

where χ_T is the indicator function of the set T.

Received June 24, 1969.

Using Doob's device ([1] page 231) one obtains the following version of the strong law of large numbers

THEOREM 2.1. *Under Hypothesis* (D_0), *if* $f: S \to (-\infty, \infty)$ *is* $\mathscr{B}_S$*-measurable and* $P_2^s[|f|] < \infty$, *then*

$$P\{\mathbf{x} \in \mathscr{X}^{\mathbf{N}} : \lim_{n\to\infty} Q_n^{\mathbf{x}}[f] = P_2^s[f]\} = 1, \tag{2.2}$$

where $P_2^s[f] = P^s[f \circ \pi_2] = \int_S f(\zeta, \eta) p(\zeta, d\eta) p^s(d\zeta)$ *and where the superscript s indicates that the expectation is taken with respect to the stationary initial distribution* $p^s \mid \mathscr{B}$.

COROLLARY 2.1. (See the classical Glivenko–Cantelli Theorem and its generalization in [4] II, Theorem 7.1). *Under Hypothesis* (D_0)

$$P\{\mathbf{x} \in \mathscr{X}^{\mathbf{N}} : Q_n^{\mathbf{x}} \to P_2^s \quad \textit{weak-star as} \quad n \to \infty\} = 1. \tag{2.3}$$

Let $\mathfrak{Q}(S)$ be the space of all p-measures on $\mathscr{B}_S$. Then weak-star convergence is known to be metrizable by various metrics. Let us consider one of them, the "BL^*-norm"-metric β for which Dudley [2], Section 3, obtained his results on the speed of mean Glivenko–Cantelli convergence in the i.i.d. case.

3. Extension of Dudley's results. Following the lines pursued by Dudley we shall obtain that, if for some $K < \infty$ and $k > 2$, S can be covered up to a set A with $P_2^s(A) \leqq \varepsilon^{k/(k-2)}$ by at most $K\varepsilon^{-k}$ sets with diameter $\leqq 2\varepsilon$ whenever $0 < \varepsilon \leqq 1$, then for some $M = M(k, \gamma, \rho, K) < \infty$

$$P^s[\beta(Q_n^{\cdot}, P_2^s)] \leqq M \cdot n^{-1/k} \qquad \text{for all} \quad n \geqq 2,$$

where γ and ρ, $\rho < 1$ are positive constants entering via (D_0) ([1] V. 5).

LEMMA 3.1. *Under Hypothesis* (D_0), *if* $f: S \to (-\infty, \infty)$ *is* $\mathscr{B}_S$*-measurable and* $P_2^s[|f|^2] < \infty$, *then the following mixing property holds*

$$|P^s[(f \circ \pi_2^{(i)})(f \circ \pi_2^{(j)})] - (P_2^s[f])^2| \leqq 2\gamma^{\frac{1}{2}} \rho^{[j-(i+1)]/2} P_2^s[|f|^2] \tag{3.1}$$

for all i, j *with* $j > i+1$.

PROOF. Follows from [1] Lemma 7.1 page 222, taking $r = s = 2$ and $f = g$.

LEMMA 3.2. *Under Hypothesis* (D_0), *if* $f: S \to (-\infty, \infty)$ *is* $\mathscr{B}_S$*-measurable and* $P_2^s[|f|^2] < \infty$ *then an upper estimate for the variance of the random variable* $Q_n^{\cdot}[f] : \mathscr{X}^{\mathbf{N}} \to (-\infty, \infty)$ *is obtained by*

$$\sigma^2_{P^s}(Q_n^{\cdot}[f]) \leqq [3 + 4\gamma^{\frac{1}{2}}\rho^{\frac{1}{2}}/(1-\rho^{\frac{1}{2}})] \cdot P_2^s[|f|^2] \cdot n^{-1} \qquad \textit{for all} \quad n \geqq 2. \tag{3.2}$$

PROOF. Let $n \geqq 2$ be arbitrary; then

$$\begin{aligned}
\sigma^2_{P^s}(Q_n^{\cdot}[f]) &= P^s[(n^{-1}\textstyle\sum_{i=0}^{n-1} f \circ \pi_2^{(i)} - P_2^s[f])^2] \\
&= n^{-1}P_2^s[|f|^2] - (P_2^s[f])^2 + 2n^{-2}\textstyle\sum_{0 \leqq i < j}^{n-1} P^s[(f \circ \pi_2^{(i)})(f \circ \pi_2^{(j)})] \\
&= n^{-1}P_2^s[|f|^2] - (P_2^s[f])^2 + 2n^{-2}\textstyle\sum_{0 \leqq i < j-1 \leqq n-2} P^s[(f \circ \pi_2^{(i)}) \\
&\qquad \cdot (f \circ \pi_2^{(j)})] + 2n^{-2}\textstyle\sum_{i=0}^{n-2} P^s[(f \circ \pi_2^{(i)})(f \circ \pi_2^{(i+1)})].
\end{aligned}$$

Using (3.1) and Schwarz's inequality we obtain

$$\begin{aligned}\sigma^2_{P^s}(Q_n^{\cdot}[f]) &\leqq n^{-1}P_2{}^s[|f|^2]-(P_2{}^s[f])^2 \\ &\quad +2n^{-2}\textstyle\sum_{0\leqq i<j-1\leqq n-2}\{(P_2{}^s[f])^2+2\gamma^{\frac12}\rho^{[j-(i+1)]/2}\cdot P_2{}^s[|f|^2]\} \\ &\quad +2n^{-2}\textstyle\sum_{i=0}^{n-2}P_2{}^s[|f|^2] \\ &\leqq 3n^{-1}P_2{}^s[|f|^2]-(3n-2)n^{-2}(P_2{}^s[f])^2 \\ &\quad +4\gamma^{\frac12}n^{-2}P_2{}^s[|f|^2]\cdot\textstyle\sum_{0\leqq i<j-1\leqq n-2}\rho^{[j-(i+1)]/2}.\end{aligned}$$

Straightforward computations yield that

$$\textstyle\sum_{0\leqq i<j-1\leqq n-2}\rho^{[j-(i+1)]/2}=(n-2)\rho^{\frac12}/(1-\rho^{\frac12})-(\rho-\rho^{n/2})/(1-\rho^{\frac12})^2;$$

hence

$$\sigma^2_{P^s}(Q_n^{\cdot}[f])\leqq[3+4\gamma^{\frac12}\rho^{\frac12}/(1-\rho^{\frac12})]P_2{}^s[|f|^2]n^{-1}. \quad \square$$

COROLLARY 3.1. *Under Hypothesis* (D_0), *if* $S_j\in\mathscr{B}_S$, $j=1,2,\cdots,m$, *are disjoint with union* T, *then for all* $n\geqq 2$

(i) $P^s[\sum_{j=1}^m\{Q_n^{\cdot}(S_j)-P_2{}^s(S_j)\}^2]\leqq\alpha(\gamma,\rho)P_2{}^s(T)n^{-1}$ *and*

(ii) $P^s[\sum_{j=1}^m|Q_n^{\cdot}(S_j)-P_2{}^s(S_j)|]\leqq[\alpha(\gamma,\rho)P_2{}^s(T)n^{-1}]^{\frac12}\cdot m^{\frac12}$, *where*

$$\alpha(\gamma,\rho):=3+4\gamma^{\frac12}\rho^{\frac12}/(1-\rho^{\frac12}). \tag{3.3}$$

PROOF. (ii) follows from (i) using Schwarz's inequality. To prove (i) we apply Lemma 3.2 with $f=\chi_{S_j}$, $j=1,2,\cdots,m$ and obtain

$$\begin{aligned}P^s[\textstyle\sum_{j=1}^m\{Q_n^{\cdot}(S_j)-P_2{}^s(S_j)\}^2]&=\textstyle\sum_{j=1}^m\sigma^2_{P^s}(Q_n^{\cdot}(S_j))\leqq\alpha(\gamma,\rho)\sum_{j=1}^m P_2{}^s(S_j)n^{-1}\\ &=\alpha(\gamma,\rho)P_2{}^s(T)n^{-1}.\end{aligned}$$

Let $BL=BL(S)$ be the Banach space of all bounded Lipschitzian real-valued functions f on S with the norm

$$||f||_{BL}:=||f||_\infty+\sup_{y\neq z}|f(y)-f(z)|/d(y,z),$$

where d denotes the metric in S. If μ, $\nu\in\mathfrak{Q}(S)$, let $||\mu||^*_{BL}:=\sup_{||f||_{BL}\leqq 1}|\mu[f]|$ and

$$\beta(\mu,\nu):=||\mu-\nu||^*_{BL}. \tag{3.4}$$

Furthermore, given $\mu\in\mathfrak{Q}(S)$, ε, $\delta>0$, let $N(\mu,\varepsilon,\delta)$ be the minimal number of sets of diameter $\leqq 2\varepsilon$ which cover S except for a set A with $\mu(A)<\delta$.

THEOREM 3.1. *Under Hypothesis* (D_0), *suppose that for some real number* $k>2$, *there is a* $K<\infty$ *such that*

$$N(P_2{}^s,\varepsilon,\varepsilon^{k/(k-2)})\leqq K\varepsilon^{-k} \tag{3.5}$$

whenever $0<\varepsilon\leqq 1$. *Then there is an* $M=M(k,\gamma,\rho,K)<\infty$ *such that*

$$P^s[\beta(Q_n^{\cdot},P_2{}^s)]\leqq M\cdot n^{-1/k} \qquad \textit{for all} \quad n\geqq 2. \tag{3.6}$$

PROOF. For each positive integer r, S is the disjoint union of sets $S_{rj} \in \mathscr{B}_S$, $j = 0, \cdots, m_r$, where $m_r \leqq K \cdot 3^{k(r+2)}$, for $j \geqq 1$ the diameter of S_{rj} is at most 3^{-r-1}, and $P_2^s(S_{r0}) \leqq 3^{-k(r+2)/(k-2)}$. Given a positive integer $n \geqq 2$, let $\varepsilon = n^{-1/k}$. Let t be the smallest integer such that $3^{-t} < \varepsilon$ and u be the smallest integer such that $3^{-u} < \varepsilon^{(k-2)/k}$. Then

$$3^t \leqq 3/\varepsilon \quad \text{and} \quad 3^u \leqq 3\varepsilon^{(2-k)/k} \quad \text{and} \quad u \leqq t. \tag{3.7}$$

Now, in exactly the same way as in the proof of Dudley's Theorem 3.2 one obtains that $\beta(Q_n^{\cdot}, P_2^s) \leqq \varepsilon + M_u + \sum_{r=u}^{t} \{(Q_n^{\cdot} + P_2^s)(S_{r0}) + 3^{1-r} M_r\}$, where $M_r := \sum_{j=1}^{m_r} |Q_n^{\cdot}(A_{rj}) - P_2^s(A_{rj})|$, the $A_{rj} \in \mathscr{B}_S, j = 1, \cdots, m_r$, being disjoint. Thus an application of Corollary 3.1 (ii) yields

$$\begin{aligned} P^s[\beta(Q_n^{\cdot}, P_2^s)] &\leqq \varepsilon + (\alpha(\gamma, \rho) m_u/n)^{\frac{1}{2}} + \sum_{r=u}^{t} \{2P_2^s(S_{r0}) + 3^{1-r}(\alpha(\gamma, \rho) m_r/n)^{\frac{1}{2}}\} \\ &\leqq \varepsilon + (\alpha(\gamma, \rho)K/n)^{\frac{1}{2}} \cdot 3^{k(u+2)/2} + \sum_{r=u}^{t} \{2 \cdot 3^{-k(r+2)/(k-2)} \\ &\qquad + 3^{1-r}(3^{k(r+2)}\alpha(\gamma, \rho)K/n)^{\frac{1}{2}}\} \\ &\leqq \varepsilon + (\alpha(\gamma, \rho)K/n)^{\frac{1}{2}} \cdot 3^{k(u+2)/2} + 2 \cdot 3^{-k(u+2)/(k-2)}[(1 - 3^{-k(t+1-u)/(k-2)})/ \\ &\qquad (1 - 3^{-k/(k-2)})] \\ &\qquad + (\alpha(\gamma, \rho)K/n)^{\frac{1}{2}} \cdot 3^{k+1} 3^{(k-2)u/2}[(1 - 3^{(k-2)(t+1-u)/2})/(1 - 3^{(k-2)/2})] \\ &\leqq \varepsilon + (\alpha(\gamma, \rho)K/n)^{\frac{1}{2}}\{3^{k(u+2)/2} + 27 \cdot 3^{(k-2)(t+3)/2}[1 - 3^{-(k-2)(t+1-u)/2}]/ \\ &\qquad (3^{(k-2)/2} - 1)\} \\ &\qquad + 2 \cdot 3^{-k(u+2)/(k-2)}/(1 - 3^{-k/(k-2)}). \end{aligned}$$

Using (3.7) and the fact that $1 - 3^{-(k-2)(t+1-u)/2} \leqq 1$ we obtain

$$\begin{aligned} P^s[\beta(Q_n^{\cdot}, P_2^s)] &\leqq \varepsilon + (\alpha(\gamma, \rho)K/n)^{\frac{1}{2}} \cdot \varepsilon\{3^{3k/2}\varepsilon^{-k/2} + 27 \cdot (3^{2(k-2)}\varepsilon^{-k/2})/3^{(k-2)/2} - 1)\} \\ &\qquad + \varepsilon\{2 \cdot 3^{3k/(2-k)}/(1 - 3^{-k/(k-2)})\}, \end{aligned}$$

hence $n^{-1/2}\varepsilon^{-k/2} = 1$ yields the result with $M := 1 + (\alpha(\gamma, \rho)K)^{\frac{1}{2}}\{3^{3k/2} + 27 \cdot (3^{2(k-2)})/(3^{(k-2)/2} - 1)\} + 2 \cdot 3^{3k/(2-k)}/(1 - 3^{-k/(k-2)})$, where $\alpha(\gamma, \rho)$ is defined by (3.3).

Let us remark that Dudley's Corollary 3.3 as well as his Proposition 3.4 on page 43 also hold in the present case. The same is true for Theorem 4.1 in [2] page 44, concerning ρ-convergence. substituting Dudley's argumentation via his Proposition 3.1 by Corollary 3.1 (ii) of the present paper.

Finally, according to the definition (3.4) of the metric β in $\mathfrak{Q}(S)$, Theorem 3.1 yields the following Corollary.

COROLLARY 3.2. *Under Hypothesis* (D_0), *suppose that for some real number* $k > 2$, *there is a* $K < \infty$ *such that* (3.5) *holds whenever* $0 < \varepsilon \leqq 1$. *Then there is an* $M = M(k, \gamma, \rho, K) < \infty$ *such that for every* $a > 0$ *and each subfamily* $UBL \subset BL$ *with* $\sup_{f \in UBL} ||f||_{BL} =: C < \infty$

$$(P^s)^*\{\mathbf{x} \in \mathscr{X}^{\mathbf{N}} : \sup_{f \in UBL} |Q_n^{\mathbf{x}}[f] - P_2^s[f]| > a\} \leqq MCa^{-1}n^{-1/k}$$

for all $n \geqq 2$, *where* $(P^s)^*$ *denotes the outer measure pertaining to* P^s.

Acknowledgment. The author wishes to thank the referee for his detailed comments.

REFERENCES

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

[2] DUDLEY, R. M. (1969). The speed of mean Gliwenko–Cantelli convergence. *Ann. Math. Statist.* **40** 40–50.

[3] GÄNSSLER, P. On the existence and speed of convergence of consistent asymptotically normal estimators for Markov processes. In preparation.

[4] PARTHASARATHY, K. R. (1967). *Probability Measures on Metric Spaces*. Academic Press, New York.