*The Annals of Probability*
1973, Vol. 1, No. 2, 362-364

# A COUNTEREXAMPLE TO PEREZ'S GENERALIZATION OF THE SHANNON-MCMILLAN THEOREM

BY J. C. KIEFFER

*University of Missouri-Rolla*

A counterexample is given to a result of Perez which makes a statement about the convergence of a sequence of logarithms of Radon-Nikodym derivatives. The result, if true, would have been a generalization of the Shannon-McMillan theorem of information theory.

Perez ([1], Theorem 2.3 and Corollary 3.3) gives a result which is a generalization of the Shannon-McMillan theorem of information theory. It is the purpose of this note to show that Perez's result is false by providing a simple counterexample. We first state Perez's result and then proceed with the formulation of the counterexample.

*Statement of Perez's result.* Let $P$, $Q$ be probability measures on a measurable space $(\Omega, \mathscr{F})$. Let $X_1, X_2, X_3, \cdots$ be a sequence of measurable maps from this space to another. We further suppose that $P$ and $Q$ are stationary measures with respect to this sequence. For $n = 1, 2, 3, \cdots$, let $\mathscr{F}_n$ be the sub sigma-field of $\mathscr{F}$ generated by $X_1, X_2, \cdots, X_n$, and let $P_n(Q_n)$ be the restriction of $P(Q)$ to $\mathscr{F}_n$. We suppose for each $n$ that $P_n$ is absolutely continuous with respect to $Q_n$; we let $f_n$ denote the Radon-Nikodym derivative of $P_n$ with respect to $Q_n$. Perez's result states that if $\lim_{n\to\infty} n^{-1} \int \log f_n \, dP$ exists and is finite, then $\lim_{n\to\infty} n^{-1} \log f_n$ exists in the sense of $L^1(P)$ convergence and also in the sense of a.e. $[P]$ convergence. (All logarithms we take to the base 2.)

*Convex sequences.* To formulate the counterexample we need certain results about convex sequences. A sequence of real numbers $c_1, c_2, \cdots$, is *convex* if $c_{n+2} - 2c_{n+1} + c_n \geqq 0$, $n = 1, 2, 3, \cdots$. It is well known (see [2]) that a non-negative convex sequence $c_1, c_2, \cdots$ converging to zero satisfies

(1) $$\sum_{i=n}^{\infty} (i - n + 1)(c_{i+2} - 2c_{i+1} + c_i) = c_n, \qquad n = 1, 2, \cdots.$$

The following lemma is useful in constructing convex sequences.

LEMMA. *Let $a_1, a_2, \cdots$ be a sequence of real numbers such that*

(a) $a_n \geqq 2$, $n = 1, 2, \cdots$
(b) $a_{n+1} \geqq a_n - n^{-1}$, $n = 1, 2, \cdots$.

*Then the sequence $(2^{-(n-1)a_n})_1^\infty$, is convex.*

Received May 2, 1972.
*AMS* 1970 *subject classifications.* Primary: 94-A15, 60F99; Secondary: 28-A65.
*Key words and phrases.* Shannon-McMillan theorem, information theory, Radon-Nikodym derivatives.

PROOF. From (b) we have $1 - na_{n+1} \leqq -na_n + 2$. From (a) we have $-na_n + 2 \leqq -na_n + a_n = -(n-1)a_n$. It follows that $2^{1-na_{n+1}} \leqq 2^{-(n-1)a_n}$. The convexity condition $2^{-(n+1)a_{n+2}} - 2(2^{-na_{n+1}}) + 2^{-(n-1)a_n} \geqq 0$ is now satisfied.

We construct some sequences we will need later.

Let $a_1, a_2, \cdots$ be a sequence such that

(a) $|a_{n+1} - a_n| \leqq n^{-1}$, $n = 1, 2, \cdots$

(b) $2 \leqq a_n \leqq 3$, $n = 1, 2, \cdots$

(c) $a_n = 2$ and $a_n = 3$ for infinitely many $n$. (It is not hard to see that such a sequence exists.) Define the sequences $(p_n)_1^\infty$ and $(q_n)_1^\infty$ as follows:

$$p_n = 2^{-(n-1)a_n}, \qquad q_n = 2^{-(n-1)(5-a_n)}, \qquad n = 1, 2, \cdots .$$

Using the lemma we see that these two sequences are convex; furthermore, they are positive, converge to zero, and $p_1 = q_1 = 1$.

*The counterexample.* We take $\Omega$ to consist of all doubly infinite sequences $(\cdots, x_1, x_2, \cdots)$ that can be formed from 0, 1, 2. $\mathscr{F}$ is the usual product sigma-field. Take $X_n$ to be the $n$th coordinate mapping, $n = 1, 2, \cdots$. We take $P$ to be the discrete probability measure which assigns probability $\frac{1}{2}$ to the sequence identically equal to 0 and to the sequence identically equal to 2. $P$ is clearly stationary.

$Q$ is the discrete probability measure defined as follows:

(a) For $n = 1, 2, \cdots, Q$ assigns probability $\frac{1}{2}(p_{n+2} - 2p_{n+1} + p_n)$ to each of the $n$ periodic sequences in $\Omega$ that can be formed by repeating the block of digits consisting of a one followed by $n-1$ zeroes.

(b) For $n = 1, 2, \cdots, Q$ assigns probability $\frac{1}{2}(q_{n+2} - 2q_{n+1} + q_n)$ to each of the $n$ periodic sequences formed from repeating the block consisting of a one followed by $n-1$ twos.

Referring back to property (1) of convex sequences, we see that the probabilities sum to one. Therefore $Q$ is a probability measure; it is easily seen to be stationary.

$P_n$ is absolutely continuous with respect to $Q_n$ because

(a) $Q_n(X_1 = 0, \cdots, X_n = 0) = \frac{1}{2}\sum_{i=n+1}^{\infty}(i-n)(p_{i+2} - 2p_{i+1} + p_i) = \frac{1}{2}p_{n+1} > 0$; and

(b) $Q_n(X_1 = 2, \cdots, X_n = 2) = \frac{1}{2}q_{n+1} > 0$. (Property 1 of convex sequences was again used.)

Also,

$$\lim_{n\to\infty} \frac{1}{n} \int \log f_n \, dP = \lim_{n\to\infty} \frac{1}{2n}\left(\log \frac{1}{p_{n+1}} + \log \frac{1}{q_{n+1}}\right) = \frac{5}{2} .$$

Therefore every hypothesis of Perez's result is satisfied. However, $n^{-1}\log f_n(\cdots, 0, 0, \cdots) = n^{-1}\log p_{n+1}^{-1}$, and therefore has no limit as $n \to \infty$. Therefore the sequence of functions $(n^{-1}\log f_n)_1^\infty$ cannot converge a.e. $[P]$ or in the $L^1(P)$ sense.

*Final remark.* Suppose in Perez's result the hypothesis that $Q$ be stationary is replaced with the requirement that $X_1, X_2, \cdots$ be a Markov process with respect to $Q$, with stationary transition probabilities. Then a true theorem is obtained, which is again a generalization of the Shannon–McMillan Theorem. This theorem was proved by Moy [3] by making essential use of martingale theory and by this author [4] with no use of martingale theory. A version of the theorem for continuous time processes $(X_t: t > 0)$ can also be obtained (see [1]).

## REFERENCES

[1] Perez, A. (1964). Extensions of Shannon-McMillan's Limit Theorem to more general stochastic processes. *Transactions of the Third Prague Conference on Information Theory: Statistical Decision Functions and Random Processes.* Prague, 545–574.

[2] Katznelson, Y. (1968). *An Introduction to Harmonic Analysis.* Wiley, New York.

[3] Moy, S. C. (1961). Generalizations of Shannon-McMillan Theorem. *Pacific J. Math.* **11** 705–714.

[4] Kieffer, J. C. (1970). A generalization of the Shannon-McMillan theorem and its application to information theory. Ph. D. Dissertation, Univ. of Illinois, Urbana.

Department of Mathematics
University of Missouri
Rolla, Missouri 65401