The Annals of Probability
1974, Vol. 2, No. 1, 161-162

ON A SET OF ALMOST DETERMINISTIC k-INDEPENDENT RANDOM VARIABLES

BY A. JOFFE

McGill University and Université de Montréal

We give an elementary example, generalizing the one of Bernstein, of a set of pairwise independent random variables which are not jointly independent. It preserves, the "not enough emphasis" characteristic of Bernstein's example in which each random variable is a function of any two of them.

1. Introduction. We give an elementary construction, generalizing the one of Bernstein, of a set of pairwise independent random variables which are not jointly independent (cf. note 2 in [1]; and [3] for other examples). It preserves, the "not enough emphasis" characteristic of Bernstein's example in which each random variable is a function of any two of them. Given a set S of random variables we say that:

(a) S is a k-independent set if any subset S' of cardinality k is formed of jointly independent random variables.

(b) S is a k-deterministic set if given any subset S' of cardinality k, all the elements of S are Borel functions of those in S'.

In [2] we have shown the existence of infinite sequences of random variables having properties (a) and (b) for $k = 2$ (pairwise independent case); however our proof was not constructive. In this note we give an elementary construction which shows the existence of arbitrary large, but finite, sets of random variables having properties (a) and (b); our purpose is essentially methodological.

2. Construction. Let p be any prime. For any $k < p$ let us define $X_1, \cdots, X_k$ to be independent random variables uniformly distributed on $[0, 1, \cdots, p - 1]$, i.e.,

$$P(X_1 = i_1, \cdots, X_k = i_k) = \frac{1}{p^k}, \quad i_1, \cdots, i_k = 0, 1, \cdots, p - 1 .$$

Let:

$$X_n = X_1 \oplus nX_2 \oplus n^2X_3 \oplus \cdots \oplus n^{k-1}X_k , \qquad \text{for } n = k + 1, \cdots, p \tag{1}$$

where $\oplus$ denotes addition modulo p. We have

THEOREM. *The random variables $X_1, \cdots, X_{p+1}$ are identically distributed and form a k-independent set and a k-deterministic set.*

PROOF. Z_p, the integers modulo p, forms a field. If we are given $X_{i_1}, \cdots, X_{i_k}$

Received September 11, 1970; revised August 15, 1972.
AMS 1970 *subject classifications.* Primary 60B05; Secondary 60-01.
Key words and phrases. Pairwise independence, independence, Bernstein.

we may without loss of generality assume that $1 \leqq i_1 < \cdots < i_l \leqq k < i_{l+1} < \cdots < i_k$. (1) shows that we can solve the system of equations for $X_1, \cdots, X_k$ since the determinant of the system is essentially a Vandermonde determinant which does not vanish modulo p. Z_p being a field there is a 1—1 correspondence between $(X_{i_1}, \cdots, X_{i_k})$ and $(X_1, \cdots, X_k)$. Denote this by $(X_{i_1}, \cdots, X_{i_k}) = g_{i_1,\cdots,i_k}(X_1, \cdots, X_k)$. Since $X_n, n = 1, \cdots, p+1$ are determined by $X_1, \cdots, X_k$ the last part of the theorem is proved. Note that $g_{i_1,\cdots,i_k}$ defines an automorphism of the direct product $Z_p{}^k$ of the additive group Z_p with itself. Let us compute the distribution of $(X_{i_1}, \cdots, X_{i_k})$: for any $(j_1, \cdots, j_k) \in Z_p{}^k$ we have $P[(X_{i_1}, \cdots, X_{i_k}) = (j_1, \cdots, j_k)] = P[g_{i_1,\cdots,i_k}(X_1, \cdots, X_k) = (j_1, \cdots, j_k)] = P[(X_1, \cdots, X_k) = g^{-1}_{i_1,\cdots,i_k}(j_1, \cdots, j_k)] = p^{-k}$; $P[X_i = j_l] = \sum_{i_r \in Z_p, r \neq l} P[(X_{i_1}, \cdots, X_{i_k}) = (j_1, \cdots, j_k)] = p^{-1}$, which proves the theorem.

3. Remarks. There are many ways of modifying the above examples; for $k = 2$, proceeding as above but defining $X_n = X_{n-2} \oplus X_{n-1}$ for $n \geqq 2$, one is led into elementary arithmetic involving the Fibonacci sequence. In fact one can obtain similar examples by considering random variables which take values in any compact Abelian group G; this was suggested by the referee; see also [2] and [4], pages 113–115. For instance if $G = R/Z$ (the real numbers modulo one), one can show by using the above construction, the existence of a sequence of random variables forming a k-independent set such that any $k+1$ of them are dependent (but not k-deterministic). Adapting the above construction to our results in [2] it follows easily that infinite sequences of random variables exist which are k-independent and k-deterministic. We do not know, however, how to construct them explicitly.

Acknowledgment. We thank D. Dawson who aroused our interest in this problem, I. Connell for a helpful conversation and the referee who greatly improved our original presentation.

REFERENCES

[1] FELLER, W. (1959). Non-Markovian processes with the semigroup property. *Ann. Math. Statist.* **30** 1252–1253.

[2] JOFFE, A. (1971). On a sequence of almost deterministic pairwise independent random variables. *Proc. Amer. Math. Soc.* **29** 381–382.

[3] GEISSER, S. and MANTEL, N. (1962). Pairwise independence of jointly dependent variables. *Ann. Math. Statist.* **33** 290.

[4] LETAC, G. (1970). *Problèmes de probabilité*. P.U.F. Paris.

DEPARTMENT DE MATHÉMATIQUES
UNIVERSITÉ DE MONTREAL
CASE POSTALE 6128
MONTREAL 101, CANADA