The main achievement of the research conducted under this contract was the sharpening of techniques for studying the asymptotic behaviour of the empirical measure $P_n$ constructed from independent sampling on a distribution $P$. The best results obtained used symmetrization and combinatorial methods to bound the deviations of $P_n$ from $P$. This gave central limit theorems for $n^{1/2} (P_n - P)$, which generalized the classical Donsker theorem for empirical distribution functions [2], [4]. These were applied to determine the limiting distribution of some statistics expressible as functionals $T_nP$ of the empirical measure, as in the central limit theorem for the optimal $k$-means cluster centres [5]. It turned out that these methods also had applications to the communications problem of quantization. The invited paper [6] explores one such connection.

Gary Oehlert and Adrian Baddeley were supported as research assistants during parts of the contract period. Gary completed a Ph.D. on the asymptotic behaviour of alternative estimators of the mean; he now holds the position of Assistant Professor of Statistics at Princeton. Adrian contributed a lot of geometric know-how to the $k$-means problem. Some of our theoretical expectations of the rotational instability of these estimators are now being detected for practical quantization algorithms.

As a by-product of this work and related graduate courses, a new approach to the study of weak convergence methods has emerged. Pursuit of this approach and a widening of the applications of the empirical process tools -- to cover rank tests, estimation from censored data, and other problems in quantization -- seems the natural next step.

Publications:
**Title:** Final Report on Contract F49620-79-C-0164

**Type of Report & Period Covered:** Final - 7/1/79 to 9/30/81

**Author:** David Pollard

**Performing Organization Name and Address:**
Yale University
Dept of Statistics
New Haven, CT 06520

**Contract or Grant Number(s):** F49620-79-C-0164

**Report Date:** Nov. 30, 1981

**Number of Pages:** Two (2)

**Security Class. (of this Report):** Unclassified

**DISTRIBUTION STATEMENT (of this Report):**
Approved for public release; distribution unlimited.

**Abstract:**
A central limit theorem for empirical processes was proved. Applications to the electrical engineering problem of quantization and the statistical k-means clustering procedure were explored.

**Key Words:**
Empirical processes, central limit theorem, quantization, k-means clustering.


PROPERTY INVENTORY

No property items were purchased under this contract, so there is no property to inventory.