J. Statist. Comput. Simul., 1978, Vol. 6, pp. 257–263

Printed in Great Britain

An Efficient Approximate Algorithm for the Kolmogorov–Smirnov and Lilliefors Tests†

TEOFILO GONZALES, SARTAJ SAHNI and W. R. FRANTA
Department of Computer Science, University of Minnesota, Minneapolis, Minnesota, U.S.A.

(*Received October 28, 1976*)

In an earlier paper we presented a linear time algorithm for computing the Kolmogorov–Smirnov and Lilliefors test statistics. In this paper we present a linear time approximate algorithm which requires less memory than the previous algorithm.

KEYWORDS and PHRASES: Kolmogorov–Smirnov test, Lilliefors test, exact and approximate algorithms, time and space complexity.

CR Categories: 5.25, 5.5

1. INTRODUCTION

The Kolmogorov–Smirnov and Lilliefors tests allow us to evaluate the hypothesis that a collected data set, i.e., a random sample $X_1, \ldots, X_n$, was drawn from a specified continuous distribution function $F(x)$. For both tests, a determination is made of the numeric difference between the specified distribution function $F(X)$ and the sample distribution function (X) defined as:

$$S(X) = j/n, j = \{\text{number of points} \leqq X\}. \tag{1.1}$$

If the sample, $X_1, \ldots, X_n$, has been sorted into nondecreasing order so that $X_1 \leqq X_2 \leqq \ldots \leqq X_n$, then the Kolmogorov–Smirnov statistics $K^+_{\max}$ (maximum positive) $K^-_{\max}$ (maximum negative) and $K_{\max}$ (maximum absolute) deviations

†This research was supported in part by NSF grant DCR 74-10081.