DISTRIBUTION OF THE ANDERSON-DARLING STATISTIC

By Peter A. W. Lewis

IBM Watson Laboratory, New York[1]

In [1] and [2] Anderson and Darling proposed the use of the statistic

$$W_n^2 = n \int_{-\infty}^{\infty} \frac{[G_n(x) - G(x)]^2}{G(x)[1 - G(x)]} \, dG(x) \tag{1}$$

for testing the hypothesis that a sample of size n has been drawn from a population with a specified continuous cumulative distribution function $G(x)$. In (1) $G_n(x)$ is the empirical distribution function defined on the sample of size n.

We consider here the problem of determining and tabulating the distribution function, $F(z; n) = \Pr\{W_n^2 \leqq z\}$, of this statistic. In [1], the asymptotic distribution of this statistic under the null hypothesis was derived and, rewritten in a form convenient for computation, it is given by

$$\begin{aligned} F(z; \infty) &= \lim_{n \to \infty} \Pr\{W_n^2 \leqq z\} \\ &= \sum_{j=0}^{\infty} a_j (z b_j)^{\frac{1}{2}} \exp[-b_j/z] \int_0^{\infty} f_j(y) \exp[-y^2] \, dy, \end{aligned} \tag{2}$$

where

$$\begin{aligned} f_j(y) &= \exp[\tfrac{1}{8} z b_j / (y^2 z + b_j)], \\ a_j &= \frac{(-1)^j (2)^{\frac{1}{2}} (4j+1) \Gamma(j + \frac{1}{2})}{j!}; \qquad b_j = \tfrac{1}{8} (4j+1)^2 \pi^2. \end{aligned} \tag{3}$$

Using the calculated values of the a_j's and b_j's, and the fact that

$$\int_0^{\infty} f_j(y) e^{-y^2} \, dy \leqq \tfrac{1}{2} (\pi)^{\frac{1}{2}} \exp[z/8],$$

it can be determined that no more than two terms of the sum ($j = 0, 1$) are needed to evaluate $F(z; \infty)$ to five decimal places over the range of z which is of interest. This range is $0 \leqq z \leqq 8$, since for all n, $F(8; n) = 1.000$, rounded to three decimal places. The integral in each term of the sum was evaluated numerically using Hermite-Gauss quadrature numerical-integration formulas (p. 327 of [3], [4]). This method of numerical integration is very efficient in terms of computing time and gives sufficient accuracy to determine $F(z; \infty)$ to five decimal places.

The results of these calculations of $F(z; \infty)$, rounded to four decimal places,

Received November 23, 1960; revised June 13, 1961.

[1] Present address: IBM Research Laboratory, Monterey and Cottle Roads, San Jose 14, California.