THE KOLMOGOROV-SMIRNOV, CRAMÉR-VON MISES TESTS

D. A. DARLING[1]

University of Chicago, University of Michigan

1. Preface. This is an expository paper giving an account of the "goodness of fit" test and the "two sample" test based on the empirical distribution function—tests which were initiated by the four authors cited in the title. An attempt is made here to give a fairly complete coverage of the history, development, present status, and outstanding current problems related to these topics.

The reader is advised that the relative amount of space and emphasis allotted to the various phases of the subject does not reflect necessarily their intrinsic merit and importance, but rather the author's personal interest and familiarity. Also, for the sake of uniformity the notation of many of the writers quoted has been altered so that when referring to the original papers it will be necessary to check their nomenclature.

2. The empirical distribution function and the tests. Let $X_1, X_2, \cdots, X_n$ be independent random variables (observations) each having the same distribution function $U(x) = \Pr\{X_i < x\}$ and put

$$\epsilon(x) = \begin{cases} 1 & x \geqq 0 \\ 0 & x < 0. \end{cases} \tag{2.1}$$

Then the (random) function

$$F_n(x) = \frac{1}{n} \sum_{j=1}^{n} \epsilon(x - X_j) \tag{2.2}$$

is called the *empirical distribution function* of the data. Clearly $F_n(x)$ is the proportion of the X_i, $i = 1, 2, \cdots, n$, which are less than x.

It is easy to calculate the first and second order moments

$$\begin{aligned} E(F_n(x)) &= U(x), \\ \operatorname{Cov}(F_n(x), F_n(y)) &= E(F_n(x)F_n(y)) - U(x)U(y) \\ &= \frac{1}{n} c(U(x), U(y)), \end{aligned}$$

where

$$c(s, t) = \min(s, t) - st = \begin{cases} s(1 - t) & s \leqq t \\ t(1 - s) & s \geqq t, \end{cases} \tag{2.3}$$
$$0 \leqq s, t \leqq 1.$$

We quote a few classical consequences of the definition (2.2):

Received April 10, 1957. Special Invited Paper read before the Institute of Mathematical Statistics, August 21, 1956.

[1] Research carried out at the Statistical Research Center, University of Chicago, under sponsorship of the Statistics Branch Office of Naval Research. Reproduction in whole or in part is permitted for any purpose of the United States Government.