AD-A253 575

2

AN APPRECIATION OF KOLMOGOROV'S 1933 PAPER

BY

M. A. STEPHENS

TECHNICAL REPORT NO. 453

JUNE 15, 1992

PREPARED UNDER CONTRACT

N00014-92-J-1254 (NR-042-267)

FOR THE OFFICE OF NAVAL RESEARCH

DEPARTMENT OF STATISTICS

STANFORD UNIVERSITY

STANFORD, CALIFORNIA

AN APPRECIATION OF KOLMOGOROV'S 1933 PAPER

BY

M. A. STEPHENS

TECHNICAL REPORT NO. 453

JUNE 15, 1992

Prepared under contract

N00014-92-J-1254 (NR-042-267)

For the Office of Naval Research

Herbert Solomon, Project Director

DEPARTMENT OF STATISTICS

STANFORD UNIVERSITY

STANFORD, CALIFORNIA

AN APPRECIATION OF KOLMOGOROV'S 1933 PAPER

by M.A. Stephens

I. Introduction

In 1933, A. N. Kolmogorov (1933a) published a short but landmark paper in the Italian Actuarial Journal. He formally defined the empirical distribution function (EDF), and then enquired how close this would be to the true distribution F(x) when this is continuous. This leads naturally to the definition of what has come to be known as the Kolmogorov statistic (or sometimes the Kolmogorov-Smirnov Statistic) D, and Kolmogorov not only then demonstrates that the difference between the EDF and F(x) can be made as small as we please as the sample size n becomes larger, but also gives a method of calculating the distribution of D at specified points, for finite n, and uses this to give the asymptotic distribution of D. The ideas in this paper have formed a platform for a vast literature, both of interesting and important probability problems and, also, concerning methods of using the Kolmogorov statistic (and also other statistics) for testing fit to a distribution. This literature continues with great strength today, after over 50 years, showing no signs of diminishing. It is evident that the ideas set in motion by Kolmogorov are of paramount importance in statistical analysis, and variations on the probabilistic problems, including modern methods of treating them, continue to hold attention.

2. A. N. Kolmogorov - early years and position in 1933.

Andrei Nikolaevich Kolmogorov was born on April 25, 1903. His father was an agronomist who later died in the aftermath of the Revolution; his mother died shortly after his birth and he was brought up by his mother's sister. He was taught by his aunts until he was seven and then went to a gymnasium in Moscow, to which he later gave much credit for his early training. He was interested early on in mathematics, but also in biology and Russian history: he widened these interests even more in later life to include, for example, methods of education and poetry. He entered Moscow University in 1920 to study physics and mathematics, but continued his studies in history. He was a student during, of course, very difficult times in Russia and in 1922, to augment his income, he became a schoolteacher while still a student, a position he held for three years. Nevertheless, he quickly came to the attention of the Professors at Moscow and, as quickly, began to produce original results in various areas of mathematics - especially in set theory and Fourier Series. In 1924 he began his lifetime interest in probability theory and, in 1925, published his first paper in this field with A. Y. Khinchin. Also in 1925, Kolmogorov graduated from Moscow University and became a postgraduate student. In the next years he published fundamental work on laws of large numbers; he regarded such laws, the study of which began with Bernoulli, as the true beginnings of probability. By the time he finished as a postgraduate (as in many European countries at the time, a thesis degree was not deemed necessary), Kolmogorov had written nearly twenty mathematical papers and, in June 1929, he joined the Institute of Mathematics and Mechanics at Moscow University as a faculty member. Two years later he became Professor and two years more saw him appointed Director of the Scientific and Research Institute of Mathematics at the University. Earlier, he had begun his fundamental work in measure theory applied to probability, arising from his concern to have a rigorous axiomatic foundation for the subject. This first appeared as a paper in 1929 and then in 1933, the same year as the paper introduced here, he produced his classical monograph on the Foundations of Probability Theory, which was to prove so influential to the development of this subject. Between these two works appeared, in 1933, "On methods of analysis in Probability Theory", in which he exhibited the relationships between the theory of probability and the classical analytic methods of theoretical physics. This too was to become a seminal work in the theory of random processes.

The paper considered here thus came when Kolmogorov was thirty years old, at the height of his mathematical powers, already recognized in the Soviet Union, and increasingly becoming so outside its borders. It is a brilliant combination of his skill with classical probability arguments combined, as we shall see, with his abilities in mathematical analysis.

For the above summary, I am greatly indebted to the review of Kolmogorov's life by Shiryaev (1989); a biography of Kolmogorov is also given in Kotz, Johnson and Read (1989).

3. Summary of the paper.

In this section the contents of the paper will be outlined in more detail than that given earlier; in subsequent sections we show some of the ways in which this short article led to advances across the broad fields of probability and statistics.

Suppose a random sample is given of n values of X; these are ordered and labelled so that $X_1 \le X_2 \le \ldots \le X_n$. In more modern notation this would be written $X_{(1)} \le X_{(2)} \le \ldots \le X_{(n)}$, but Kolmogorov's original will be used here.

(a) The function $F_n(x)$, called the empirical distribution function (EDF) is defined as

$$
\begin{array}{lll}
F_n(x) = 0 & x < X_1; & \\
F_n(x) = \dfrac{k}{n} & X_k \le x < X_{k+1} & k=1,2,\ldots \\
 & & n-1; \\
F_n(x) = 1 & X_n \le x &
\end{array}
$$

(b) Kolmogorov states that we are "almost naturally" led to ask if $F_n(x)$ is approximately equal to $F(x)$ when n assumes a very large value, and refers to von Mises' (1931) book which, only two years earlier, had introduced another statistic to measure how close $F_n(x)$ is to $F(x)$. Kolmogorov defines

$$D = \sup_x | F_n(x) - F(x) |$$

and points out the importance of answering whether $\Pr(D < \varepsilon)$ tends to 1 as $n \to \infty$, however small the ε.

(c) He answers the question by proving the following asymptotic result, expressed as Theorem I.

Let $\Phi(\lambda) = \Pr(D < \lambda/\sqrt{n})$; then $\Phi(\lambda)$, as $n \to \infty$ uniformly in λ, tends to

$$\Phi(\lambda) = \sum_{k=-\infty}^{\infty}(-1)^k e^{-2k^2\lambda^2}$$

for any continuous distribution function F(x). Some values of $\Phi(\lambda)$ are given for various λ; it is pointed out that, for small λ, $\Phi(\lambda)$ converges slowly, and the first term of the equivalent formula

$$\Phi(\lambda) = \frac{\sqrt{2\pi}}{\lambda}\sum_{k=1}^{\infty}\exp\left[-(2k-1)^2\,\pi^2/(8\lambda^2)\right]$$

then gives excellent results for $\lambda < 0.6$.

(d) The proof of the Theorem first involves the probability integral transformation $Y = F(X)$, showing that the distribution of Y is $F(y) = y$, $0 \leq y \leq 1$, namely the uniform distribution. Also, if D_y is calculated from the EDF of the Y-values given by $Y_i = F(X_j)$, $i = 1, 2, \ldots n$, then D_y will equal D. Thus the result required may be deduced assuming that the original values have a uniform distribution between 0 and 1, which we shall write U(0,1).

(e) The calculations are based on the following argument. Suppose lines U ($y = x + d$) and L ($y=x-d$) are drawn parallel to $y=F(x)=x$. For $D < d$, all the "corners" of $F_n(x)$ must lie between U and L. Suppose P_{ik} is the probability that E_{ik} occurs: E_{ik} is the event that $F_n(x)$ lies between U and L at the values $x = j/n$, for all $j \leq k$, while also, at $x=k/n$, $|F_n(k/n) - (k/n)| = i/n$. Clearly $P(\dot{D} < d)$ is then P_{0n}. Kolmogorov gives a formula for P_{ik^*}, where $k^*=k+1$, as a linear combination of the P_{ij} for $j \leq k$; the coefficients in the expression are conditional probabilities $Q_{ji}(k)$ that E_{ik^*} occurs given that E_{jk} has occurred. These linear equations can be solved for P_{ik} and, hence, for the required P_{0n}.

For practical calculations, Kolmogorov defines new quantities R_{ik} as functions of the P_{ik}; these enable R_{ik^*} to be expressed as linear combinations of R_{jk}, similar to the equations for P_{ik^*}, but with easier coefficients.

(f) At this point Kolmogorov's analytic skills are brought to bear. A Theorem II is given, describing the behavior of a random walk with steps Y_j which are integral multiples of a constant ε.

Suppose $S_k = \sum_{j=1}^{k} Y_i$, and let $S_n = i\varepsilon$ for some i.

Kolmogorov gives a result for $R_{\bar{i}n}$, the probability that S_k always lies between certain bounds, in terms of the Green's function of classical mathematical physics. The theorem gives the solution to a much more general problem than that discussed here; it is not proved in detail, but reference is made to an existing note and to one forthcoming. (Kolmogorov, 1933 b). For the particular problem concerning D, the Y_i are made to be Poisson variables, and $R_{\bar{i}n}$ is shown to be the same as R_{in} in paragraph (e) above.

The steps ε now approach zero, and the random walk becomes "tied down" to zero at the n-th step, thus becoming the Brownian bridge of modern notation; application of Theorem II with appropriate boundaries gives the asymptotic result given in Theorem I.

4. Contemporary work and the impact of the paper.

It seems fair to say that Kolmogorov regarded his paper as the solution of an interesting problem in probability, following his interests of the time, rather than a paper in statistical methodology. Apart from the casual remark that $F_n(x)$ should closely estimate $F(x)$ in some sense, no suggestion is made that $F_n(x)$ should be used for testing that F(x) is the distribution of x. This was, nevertheless, to become one of the major outgrowths of the article. Suggestions that $F_n(x)$ should be used for such a test were in the air at the time. Cramér (1928) had proposed expanding $F(x)$ in a type of Gram-Charlier series and then using as test statistics integrals of the type

$$I_j = \int \{\Delta_j(x)\}^2 dx \quad \text{, where } \Delta_j(x) = F_n(x) - \hat{F}_j(x),$$

and $\hat{F}_j(x)$ is the expansion of $F(x)$ up to the j-th term. The integral is over the support of x. The term $\Delta_j(x)$ can be thought of as the j-th component of the difference $F_n(x) - F(x)$, and the approach is reminiscent of Neyman's work on smooth tests, which appeared a few years later. In 1931, von Mises (1931) suggested that a test could be based on the statistic

$$\omega^2 = n\int \lambda(x)\left[F_n(x) - F(x)\right]^2 dx$$

where $\lambda(x)$ is a suitably chosen weight function. Von Mises suggested that $\lambda(x)$ should be constant, chosen so that $E(\omega^2) = 1$, and with this $\lambda(x)$ von Mises gave a computing formula for ω^2. The distribution of the criterion will vary with $F(x)$ under test (and also of course with $\lambda(x)$) even when this is completely specified; von Mises gave no distribution theory, but evaluated some variances of the criterion when the true distribution is uniform or normal.

Several years later, the Soviet mathematician and statistician Smirnov (1936, 1937) made a significant change in the definition of ω^2. This was to write

$$\omega^2 = n\int \lambda(F(x))\left[F_n(x) - F(x)\right]^2 dF(x)$$

so that the integral is with respect to $F(x)$ rather than to x. The criterion now becomes based on the values of $Z_i = F(X_i)$, which, as was seen above in paragraph 3(d), will be $U(0,1)$; it will now be distribution-free, that is, not dependent on the true $F(x)$. This version of the statistic, with $\lambda(F(x)) = 1$, has come to be known as W^2, the Cramér-von Mises statistic.

A notable achievement of Smirnov was to find the asymptotic distribution of W^2, in the form of a sum of weighted χ_1^2 variables.

Smirnov (1939 a,b) was also interested in Kolmogorov's work; he extended it to encompass one-sided tests and also two-sample tests.

Let $D^+ = \sup_x \{F_n(x) - F(x)\}$, and $D^- = \sup_x . \{F(x) - F_n(x)\}$;

these will have the same asymptotic distribution, which was found by Smirnov:

$$\lim_{n \to \infty} P(\sqrt{n}\, D^+ < \lambda) = 1 - e^{-2\lambda^2}$$

For two samples, suppose $F_n(x)$ and $G_m(x)$ are the EDF's of two independent random samples of sizes n, m respectively; define $N = mn/(m+n)$, and let

$D^+_{n,m} = \sup_x \{F_n(x) - G_m(x)\}$, $D^-_{n,m} = \sup_x \{G_m(x) - F_n(x)\}$ and
$D_{m,n} = \sup_x |F_n(x) - G_m(x)|$.

Smirnov shows that the asymptotic distribution of $\sqrt{N}\, D_{m,n}$ is the same as that of $\sqrt{n}D$ given in Kolmogorov's Theorem I.

Smirnov (1939a) also examined $V_n(\lambda)$, the number of crossings of $F_n(x)$ with the lines $F(x) \pm \lambda\sqrt{n}$, and showed that as $n \to \infty$, $P(V_n(\lambda) \le t\sqrt{n})$ converges to

$$\Theta(t,\lambda) = 1 - 2\sum_{m=0}^{\infty} \frac{(-1)^m}{m!} \frac{d^m}{dt^m}\left[t^m \exp\left\{-\frac{(t+2\lambda m + 2\lambda)^2}{2}\right\}\right]$$

He also gave a new proof of Kolmogorov's Theorem I, and tabulated the asymptotic distribution $\Phi(\lambda)$ in Smirnov (1939b); and in Smirnov (1944) he found the distribution of $\sqrt{n}D^+$. The table of $\Phi(\lambda)$ was later reproduced in English in Smirnov (1948). Statistics of the D^+, D^- and D type are often referred to as Kolmogorov-Smirnov statistics.

5. The war and afterwards.

Thus, over a period of about 10 years, the foundations were laid by a number of distinguished mathematicians of methods of testing fit to a distribution based on the EDF. To test the null hypothesis H_0 that $F(x)$, completely specified, is the true distribution of X, the statistics above may be calculated and referred to the appropriate distribution.

At this point the war intervened and much momentum in this field was certainly lost. Kolmogorov, himself, became involved in war work (he worked, for example, on artillery problems), which certainly brought him into greater contact with statistical analysis, and may account for an increasing interest in statistics itself. In 1948 he edited and wrote a preface to the Russian edition of Cramér's Mathematical Methods of Statistics; he protested the overly theoretical basis of the training of Soviet statisticians, a lament familiar enough outside the Russian borders. Perhaps, also, Kolmogorov was impressed by Cramér's opening, which gives great credit to British and American statisticians for advances in statistics, while admiring France and Russia for their excellence in probability; at any rate, in that year he spoke at a Tashkent Conference on Mathematical Statistics, on "Basic problems of Theoretical Statistics" and also enlightened the assembled statisticians on "The real meaning of the Analysis of Variance". This was to be followed, over the years, by many more contributions to the mainstream of statistics, while, of course, his other wide interests were maintained. These came to include, with the passing years, a strong interest in the teaching of both mathematics and statistics.

In the 1950's there was a surge of interest in Russia in the Kolmogorov-Smirnov statistics, particularly in the combinatoric problems associated with crossings and with two-sample statistics. Gnedenko and Korolyuk (1951) found the exact distributions of $D^+_{n,n}$ and of $D_{n,n}$, to compare two empirical distributions from independent samples both of size n: later Korolyuk (1955) found exact distribution theory when m is an integral multiple of n, m=np. By allowing $p \to \infty$, he deduced the exact distribution of D^+, and also the more difficult distribution of D itself.

Gnedenko and Rvaceva (1952) obtained the joint distribution of $D^+_{n,n}$, and $D^-_{n,n}$, and verified the asymptotic joint distribution already found by Smirnov in 1939; further results were given by Gnedenko (1952). Gnedenko and Mihalevic (1952a,b) discussed the number of crossings, when one distribution function $F_n(x)$ crosses the other $G_m(x)$. The interest spread to Hungary and across Asia to China; Renyi (1953) proposed several variations of Kolmogorov's statistic, such as $\sup_x \left|\{F_n(x) - F(x)\}/F(x)\right|$; Chang Li-Chien (1955) examined the ratio of $F_n(x)/F(x)$, closely related to Renyi's statistics, and Cheng Ping (1958) gave further results on crossings.

Meantime, in the western world also, EDF statistics were attracting attention. An elegant paper by Feller (1948) appeared, giving more accessible proofs of the results of both Kolmogorov and Smirnov: where Kolmogorov used Green's function to find the asymptotics from the equations for $P(D < c/n)$, Feller introduced generating functions for the component probabilities and then examined their limiting forms. He also gave a theorem on the asymptotic expectation of the number of crossings $V_n(\lambda)$ of $F_n(x)$ with the boundaries $F(x) \pm \lambda\sqrt{n}$. At about the same time, there were significant advances in methodology. Doob, in 1949, suggested that the asymptotic behaviour of EDF statistics based on $\varepsilon_n(x) = F_n(x) - F(x)$ could be found by examining the limiting behaviour of $\sqrt{n}\,\varepsilon_n(x)$, a Gaussian process, and calculating the statistics from this limiting process. According to Khmaladze (1986), in an article presenting the 1933 paper in Kolmogorov's collected works, Kolmogorov himself put forward similar ideas in a Moscow seminar towards the end of 1948, and Smirnov (1949) wrote a brief paper on the asymptotics of the Cramér-von Mises statistic. These ideas, those of Doob made rigorous by Donsker (1952), laid the foundation for a great deal of later work on the asymptotics of EDF statistics. Anderson and Darling (1952) used them to examine such statistics and introduced the statistic A^2, for which the weight function in Smirnov's version of ω^2 is $1/[F(x)\{1-F(x)\}]$. This compensates for the fact that $\varepsilon_n(x)$ must necessarily become small in the tails, by essentially dividing by the variance of $\varepsilon_n(x)$, and gives due weight to tail observations.

These developments demonstrated elegant techniques of combinatorics and analysis in the field of probability but, apart from some asymptotic tables, the practical statistician was largely neglected. However, in the 1950's, other authors were filling the gap. Massey (1950, 1951a), and Birnbaum and Tingey (1951), using new formulas and difference equations, gave tables of percentage points and of probabilities for finite sample size n, for D and D^+; these were later augmented by Miller (1956). Birnbaum (1952), using the original techniques of Kolmogorov himself, gave complete tables of the distribution of D, and a table of percentage points for n up to 100. Thus, at last - nearly twenty years after the statistic was suggested! - practical formulas and tables were available to make D available to test that F(x) is a completely specified continuous distribution.

Many years later again, Stephens (1970) used these tables to derive a modification of D. This is a simple expression in D and n which gives D^*; this is to be compared, for testing purposes, with the asymptotic points for $\sqrt{n}D$ given by Kolmogorov's Theorem I. The test is thus made easy to use without extensive tables of points for every n. Stephens (1970) also found similar modifications for D^+ and D^-, for $V = D^+ + D^-$ (see Section 7 below) and for the Cramér-von Mises W^2.

For two samples, Massey (1951b) and Drion (1952) gave tables for $D_{n,n}$ and Massey (1952) for $D_{m,n}$, mostly for n=mp, where p is an integer; practical formulas for the calculation of these statistics also began to appear in the literature. It was also pointed out (Wald and Wolfowitz (1939), Massey (1950), Birnbaum and Tingey (1951)) that D can be used to give a confidence interval for F(x), and D^+ a one-sided interval.

At this point all attempt will be abandoned to survey exhaustively the enormous literature which has grown up on Kolmogorov-Smirnov statistics and on other EDF statistics; many more properties of D, D^+ and D^- have been discovered, new methods of computing distributions have been proposed, and variants of the basic statistics have been suggested. Durbin (1973) provides a comprehensive and unifying account of developments up to that time, with many references; Niederhausen (1981b) also has references and brings together many of the computational procedures. A survey of goodness-of-fit tests is in Kendall and Stuart (1979, Vol. 2, Chap. 30) and another was given by Sahler (1968).

6. The problem of unknown parameters.

Despite the interest of mathematical statisticians, and the availability of tables, it has taken many years for the Kolmogorov-Smirnov statistics, and other EDF statistics, to become part of the regular arsenal of applied statisticians. No doubt this is because major new problems are presented if tests are to be made on $F(x)$, which we now call $F(x;\theta)$, when $F(x;\theta)$ is a continuous distribution containing parameters which are components of the vector θ, and when one or more of these components must be estimated from the given data set. For the well-established Pearson X^2 test, provided the estimation of parameters is done correctly – but how often it is not! – the asymptotic χ^2 distribution on H_0 merely changes its degrees of freedom, but for D^+, D^- and D, (and for other EDF statistics) the distribution theory will depend on the particular $F(x;\theta)$ being tested. This is so even when the unknown components of θ are estimated by maximum likelihood or another efficient method; the distributions, even asymptotic, are now stochastically much smaller than for the case when $F(x;\theta)$ is completely known. For Kolmogorov-Smirnov statistics, they depend asymptotically on the distribution of the maximum of a Gaussian process with mean zero, tied down at 0 and 1; even though the covariance can be found, this distribution remains unknown and the early techniques of Kolmogorov will not find it. The discovery of the asymptotics of D^+, D^- and D, when parameters must be estimated, thus remains a major theoretical problem in the area of Kolmogorov-Smirnov statistics.

If the unknown components of θ are only location or scale parameters, however, the distribution theory of all EDF statistics, even for finite n, will depend only on the family tested, and not on the true values of these parameters, a fact early recognized by David and Johnson (1948). In these circumstances, Durbin (1973, 1975) has shown how exact distributions of D^+ and D can be calculated for the exponential distribution $F(x;\theta) = 1 - \exp(-x/\theta)$, $x \leq 0$, with unknown scale θ, and has provided points for test purposes; for other distributions, including the normal, extreme-value, Weibull, and logistic distributions, several authors have produced Monte Carlo tables. For Cramér -von Mises statistics, the situation is different; asymptotic distributions can be found (see, e.g., Darling, 1955, Durbin, 1973, Stephens 1976) and percentage points for finite n converge rapidly to the asymptotic points. Also, for some important distributions with shape parameters, for example, the von Mises and Gamma, the asymptotic points for Cramér-von Mises statistics do not depend strongly on the true value of the shape, and a test using the estimated shape can be used (Lockhart and Stephens, 1985 a,b). The tests described above, for parameters known or unknown. have been collected together in Stephens (1986).

7. Further developments

We conclude this introduction by giving only a brief summary of some of the more important developments of Kolmogorov-Smirnov tests, with references either to basic introductory sources or to articles which themselves survey the particular area and give references.

Kolmogorov-Smirnov tests have been developed for use with right-or left-censored data (or both): these mostly use D, but some variations of Renyi-type, such as taking the supremum of $F_n(x) - F(x)$ over a restricted range of $F(x)$ or of $F_n(x)$ have also been suggested. Randomly censored data is an important problem, for example, with survival data: tests with such data often use the Kaplan-Meier estimate of $F(x)$. Hall and Wellner (1980) give a review and show how confidence bounds for the distribution can be found. A recent technique for censored data is given by Guilbaud (1988).

The statistic $V = D^+ + D^-$ has been proposed (Kuiper, 1960) for use with data on a circle, because the value of V, in contrast to those of

D^+, D^- or D, does not depend on the choice of origin. Of course V can also be used for data on a line. Pettitt and Stephens (1977) produced tables for D for the uniform distribution for discrete data, and Niederhausen (1981a) for a variance-weighted D, similar to A^2. A test for symmetry of a distribution was proposed by Smirnov (1947) and has since been extended; Gibbons (1983) gives a review of such tests. Tables for some of the above tests, and further discussion and references, are in Stephens (1983, 1986). An interesting area for future work is to provide tests for multivariate distributions.

Statistics closely related to D^+, D^- and D were proposed by Pyke (1959). Suppose x_i, $i=1,...,n$ are the order statistics for a sample from the uniform distribution; C^+ is $\max_i (x_i - i/(n+1))$, $C^- = \max_i (i/(n+1) - x_i)$, and $C = \max(C^+, C^-)$. These arise naturally in examining the Poisson process, or the periodogram in time series analysis: they are discussed by Durbin (1973).

8. Power

In terms of power, Kolmogorov-Smirnov tests tend to fall between the Pearson X^2 and the Cramér-von Mises tests. On the one hand, this might be expected, since X^2 loses information in a test for a continuous distribution by grouping the data into cells. Kac, Kiefer and Wolfowitz (1955) showed that if equi-probable cells are used for X^2, and if $\Delta = \sup_x|F_1(x) - F(x)|$ where $F_1(x)$ is the true distribution and $F(x)$ the tested distribution, D requires $n^{4/5}$ observations compared with n observations for X^2 to attain the same power for a given Δ, for large n. Thus in these circumstances, X^2 will have asymptotic relative efficiency equal to zero compared with D. Many Monte Carlo studies have confirmed this superiority of D over X^2 in most situations, especially with small samples.

On the other hand, Cramér-von Mises statistics might well be expected to be superior to D, since they make a comparison of $F_n(x)$ with $F(x)$ all along the range of x, rather than looking for a marked difference at one point. If the alternative is directional, that is, if $F_1(x) - F(x)$ is mostly positive or mostly negative, the one-sided D^+ or D^- can be very powerful. Of all the different families of goodness-of-fit statistics, Cramér-von Mises statistics provide overall powerful tests. (Stephens, 1974: see also Kendall and Stuart (1979) and Stephens (1986) for more discussion.)

9. Concluding Remarks.

If the remarks above on power may appear to weaken the appeal of D and its related statistics, it should nonetheless be emphasized that they are preferable to the much used X^2 statistic. They also have the value that they can be used, by simply adding a constant to $F_n(x)$, and subtracting it from $F_n(x)$, to give a confidence interval for $F(x)$ – an attraction in today's world where graphical display is increasingly available.

The final assessment of the article by Kolmogorov must be based not only on the elegance and power of the paper itself, but also on the pioneering role it has played in the development of statistics in the succeeding 50 years and more. It launched seriously the use of the EDF $F_n(x)$ as an estimator of $F(x)$, to be followed by its use in testing a given $F(x)$; it was the first article to give a statistic which would not depend (when the null hypothesis was true) on the distribution $F(x)$ tested; it was also the first to introduce a statistic whose asymptotic distribution could be found and easily tabulated. Kolmogorov also gave the essential technique to find the distribution for finite samples. More than 50 years later, interest in Kolmogorov's and other EDF statistics continues unabated. It is fitting, in conclusion, to note the resurgence of $F_n(x)$ in the wide use of the bootstrap: this technique, making use of the power which modern computers provide, is based on the use of $F_n(x)$ to estimate $F(x)$, just as was proposed by Kolmogorov in 1933.

This article was written to introduce Kolmogorov's paper in a forthcoming volume on the most influential articles in Statistics, to be edited by S. Kotz and N.L. Johnson.

10. References

ANDERSON, T.W. AND DARLING, D.A. (1952): Asymptotic theory of certain Goodness of Fit criteria based on stochastic processes. Ann. Math. Statist., 23, 193-212.

BIRNBAUM, Z.W. AND TINGEY, F.H. (1951): One-sided confidence contours for distribution functions. Ann. Math. Statist., 22, 592-596.

BIRNBAUM, Z.W. (1952): Numerical tabulation of the distribution of Kolmogorov's statistic for finite sample size, J. Amer. Statist. Assoc., 47., 425-441.

CRAMER, H. (1928): On the composition of elementary errors. Second paper: Statistical Applications. Skand. Aktuarietidskr, 171-180.

CHANG LI-CHIEN (1955): On the ratio of an empirical distribution function to the theoretical distribution function. Acta Math. Sinica 5, 347-368. Selected Translations in Mathematical Statistics and Probability, (1963), 4, 17-38.

CHENG PING (1958): Non-negative jump points of an empirical distribution function relative to a theoretical distribution function. Acta Math Sinica 8, 333-347. Selected Translations in Mathematical Statistics and Probability, (1962), 3, 205-224.

DARLING, D.A. (1955): The Cramer-Smirnov test in the parametric case. Ann. Math. Statist., 26, 1-20.

DAVID, F.N. AND JOHNSON, N.L. (1948): The probability integral transformation when parameters are estimated from the sample. Biometrika, 35, 182-192.

DOOB, J.L. (1949): Heuristic approach to the Kolmogorov-Smirnov Theorems. Ann. Math. Statist. 20, 393-403.

DRION, E.F. (1952): Some distribution-free tests for the difference between two empirical cumulative distribution functions. Ann. Math. Statist., 23, 563-574.

DURBIN, J. (1973): Distribution Theory for Tests Based on the Sample Distribution Function. CBMS-NSF Regional Conference Series in Applied Mathematics. Society for Industrial and Applied Math.; Philadelpnia.

DURBIN, J. (1975): Kolmogorov-Smirnov tests when parameters are estimated with applications to tests of exponentiality and tests on spacings. Biometrika, 62, 1, 5-22.

FELLER, W. (1948): On the Kolmogorov-Smirnov limit theorems for empirical distributions. Ann. Math. Statist., 19, 177-189.

GIBBONS, J.D. (1983): Kolmogorov-Smirnov symmetry test. Encyclopedia of Statistical Sciences (S. Kotz, N.L. Johnson and C.B. Read, eds.), 4, 396-398.

GNEDENKO, B.V. (1952): Some results on the maximum discrepancy between two empirical distributions. Dokl. Akad. Nauk SSSR, 82, 661-663. Selected Translations in Mathematical Statistics and Probability, (1961), 1, 73-76.

GNEDENKO, B.V. AND KOROLYUK, V.S. (1951): On the maximum discrepancy between two empirical distributions. Dokl. Akad. Nauk SSSR, 80, 525-528. Selected Translations in Mathematical Statistics and Probability, (1961), 1, 13-22.

GNEDENKO, B.V. AND MIHALEVIC, V.S. (1952a): On the distribution of the number of excesses of one empirical distribution function over another Dokl. Akad. Nauk SSSR, 82, 841-843. Selected Translations in Mathematical Statistics and Probability, (1961), 1, 83-85.

GNEDENKO, B.V. AND MIHALEVIC, V.S. (1952b): Two theorems on behaviour of empirical distribution functions. Dokl. Akad. Nauk SSSR, 85, 25-27. Selected Translations in Mathematical Statistics and Probability, (1961), 1, 55-58.

GNEDENKO, B.V. AND RVACEVA, E.L. (1952): On a problem of the comparison of two empirical distributions. Dokl. Akad. Nauk SSSR, 82. 513-516. Selected Translations in Mathematical Statistics and Probability, (1961), 1, 69-72.

GUILBAUD, O. (1988): Exact Kolmogorov-type tests for left truncated and/or right-censored data. J. Amer. Statist. Assoc., 83, 213-221.

HALL, W.J. AND WELLNER, J.A. (1980): Confidence band for a survival curve from censored data. Biometrika, 67, 133-143.

KAC, M., KIEFER, J. AND WOLFOWITZ, J. (1955): On tests of normality and other tests of goodness of fit based on distance methods. Ann. Math. Statist., 26, 189-211.

KENDALL, M.G. AND STUART, A. (1979): The Advanced Theory of Statistics. 4th edition. McMillan: New York.

KOLMOGOROV, A. (1931): Uber die analytischen Methoden in der Wahrscheinlichkeitsrechnung. Math. Ann., 104, 415-458.

KOLMOGOROV, A. (1933a): Sulla determinazione empirica di una legge di distribuzione. Inst. Ital. Attuari, Giorn., 4, 1-11.

KOLMOGOROV, A. (1933b): Uber die Grenzwertsatze der Wahrscheinlichkeitsrechnung. Bulletin (Izvestija) Academie des Sciences URSS, 363-372.

KOROLYUK, V.S. (1955): On the discrepancy of empiric distributions for the case of two independent samples. Izv. Akad. Nauk SSSR Ser. Mat. 19, 91-96. Selected Translations in Mathematical Statistics and Probability, (1963), 4, 105-122.

KOTZ, S., JOHNSON, N.L. AND READ (1989): A.N. Kolmogorov Encyclopedia of Statisticial Sciences, **10**, .

KUIPER, N.H. (1960): Tests concerning random points on a circle. Proc. Koninkl. Neder. Akad. van. Wetenschappen, A, 63, 38-47.

LOCKHART, R.A. AND STEPHENS, M.A. (1985a): Goodness-of-fit for the von Mises distribution. Biometrika, 72, 647-652.

LOCKHART, R.A. AND STEPHENS, M.A. (1985b): Goodness-of-fit tests for the gamma distribution. Technical Report, Department of Mathematics and Statistics, Simon Fraser University.

MASSEY, F.J. (1950): A note on the estimation of a distribution function by confidence limits. Ann. Math. Statist., 21, 116-119.

MASSEY, F.J. (1951a): The Kolmogorov-Smirnov tests for goodness of fit. J. Amer. Statist. Assoc., 46, 68-78.

MASSEY, F.J. (1951b): The distribution of the maximum deviation between two sample cumulative step functions. Ann. Math. Statist., 22, 125-128.

MASSEY, F.J. (1952): Distribution table for the deviation between sample cumulatives. Ann. Math. Statist., 23, 435-441.

MILLER, L.H. (1956): Table of percentage points of Kolmogorov statistics. J. Amer. Statist. Assoc., 51, 111-121.

NIEDERHAUSEN, H. (1981a): Tables of significant points for the variance-weighted Kolmogorov-Smirnov statistics. Technical report, Department of Statistics, Stanford University.

NIEDERHAUSEN, H. (1981b): Sheffer polynomials for computing exact Kolmogorov-Smirnov and Renyi type distributions. Annals of Statistics, 5, 923-944.

PETTITT, A.N. AND STEPHENS, M.A. (1977): The Kolmogorov-Smirnov Goodness-of-fit statistics with discrete and grouped data. Technometrics, 19, 205-210.

PYKE, R. (1959): The supremum and infimum of the Poisson process. Ann. Math. Statist., 30, 568-576.

RENYI, A. (1953): On the theory of order statistics. Acta Math. Acad. Sci. Hungar. 4, 191-231.

SAHLER, W. (1968): A survey of distribution-free statistics based on distances between distribution functions. Metrika, 13, 149-169.

SHIRYAEV, A.N. (1989): Kolmogorov -- Life and creative activities. Annals of Probability, 17, 866-944.

SMIRNOV, N.V. (1936,1937): Sur la distribution de w2 (criterium de M.R. v. Mises). Com. Rend. Acad. Sci. (Paris), 202 (1936), 449-452, and a paper of the same title in Russian Recueil Math. 2 (1937), 973-993.

SMIRNOV, N.V. (1939a): Ob uklonenijah empiriceskoi krivoi raspredelenija. Recueil Mathematique (Mathematiceskii Sbornik, N.S., 6(48), 13-26.

SMIRNOV, N.V (1939b) : On the estimation of the discrepancy between empirical curves of distributions for two independent samples. Bulletin mathematique de l'Universite de Moscou, 2 (1939), 2.

SMIRNOV, N.V. (1944): Approximate laws of distribution of random variables from empirical data. Uspehi Mat. Nauk, 10, 179-206.

SMIRNOV, N.V. (1947): Sur un critere de symetrie de la loi de distribution d'une variable aleatorie. Akad. Nauk SSSR, C.R. (Dokladi), 56, 11-14.

SMIRNOV, N.V. (1948): Table for estimating the goodness of fit of empirical distributions. Ann. Math. Statist., 19, 279-281.

SMIRNOV, N.V. (1949): On the Cramér-von Mises criterion (In Russian), Uspehi. Mat. Nauk, 14, 196-197.

STEPHENS, M.A. (1970): Use of the Kolmogorov-Smirnov, Cramer-von Mises and related statistics without extensive tables. J. Roy. Statist. Soc., B, 32, 115-122.

STEPHENS, M.A. (1974): EDF Statistics for goodness-of-fit and some comparisons. J. Amer. Statist. Assoc., 69, 730-737.

STEPHENS, M.A. (1976): Asymptotic results for goodness-of-fit statistics with unknown parameters. Ann. Statist., 4, 357-369.

STEPHENS, M.A. (1983): Kolmogorov-Smirnov statistics, Kolmogorov-Smirnov tests of fit. Encyclopedia of Statistical Sciences. (S. Kotz, N.L. Johnson and C.B. Read, eds.), 4, 393-396, 398-402.

STEPHENS, M.A. (1986): Tests based on EDF statistics. Chapter 4 in Goodness-of-fit Techniques (R.B. D'Agostino and M.A. Stephens, eds.). Marcel Dekker: New York.

VON MISES, R. (1931): Vorlesungen aus dem Gebiete der Angewandten Mathematik, 1, Wahrscheinlichkeitsrechnung und ihre Anwendung in der Statistik und theoretischen Physik. Leipzig and Wien.

WALD, A. AND WOLFOWITZ, J. (1939): Confidence limits for continuous distribution functions. Ann. Math. Statist., 10, 105-118.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 453	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) An Appreciation of Kolmogorov's 1933 Paper		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) M. A. Stephens		8. CONTRACT OR GRANT NUMBER(s) N0025-J-1264
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics Stanford University Stanford, CA 94305		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR-042-267
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Statistics & Probability Program Code 111		12. REPORT DATE June 15, 1992
		13. NUMBER OF PAGES 19
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Kolmogorov statistics, goodness-of-fit

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

An account is given of the paper published in 1933, in which A. N. Kolmogorov introduced the statistic which bears his name. The paper was seminal in the development of goodness-of-fit procedures. Kolmogorov's early career to that date is reviewed, and a summary is given of the extensive work which has followed from the 1933 paper.