The Annals of Statistics
1977, Vol. 5, No. 1, 88-97

THE EMPIRICAL CHARACTERISTIC FUNCTION AND ITS APPLICATIONS

BY ANDREY FEUERVERGER[1] AND ROMAN A. MUREIKA

University of Alberta and University of Toronto;
University of Alberta

Certain probability properties of $c_n(t)$, the empirical characteristic function (ecf) are investigated. More specifically it is shown under some general restrictions that $c_n(t)$ converges uniformly almost surely to the population characteristic function $c(t)$. The weak convergence of $n^{\frac{1}{2}}(c_n(t) - c(t))$ to a Gaussian complex process is proved. It is suggested that the ecf may be a useful tool in numerous statistical problems. Application of these ideas is illustrated with reference to testing for symmetry about the origin: the statistic $\int [\mathrm{Im}\, c_n(t)]^2\, dG(t)$ is proposed and its asymptotic distribution evaluated.

1. Introduction. Throughout this paper $X, X_1, X_2, X_3, \cdots$ represent independent random variables with distribution function $F(x)$ and characteristic function

$$c(t) = \int_{-\infty}^{\infty} e^{itx}\, dF(x)\ .$$

Many proposed statistical procedures for sequences of i.i.d. random variables may loosely be thought of as based on the empirical cdf $F_n(x) = N(x)/n$, where $N(x)$ is the number of $X_j \leqq x$ with $1 \leqq j \leqq n$, for example, procedures based on statistics of a Kolmogorov–Smirnov type. In view of the one-to-one correspondence between distribution functions and characteristic functions, it seems natural to investigate procedures based on the ecf defined as

$$c_n(t) = \int e^{itx}\, dF_n(x) = \frac{1}{n} \sum_{j=1}^{n} e^{itX_j}\ .$$

The property, for example, that a characteristic function is real if and only if the corresponding distribution function is symmetric about the origin, suggests looking at statistics such as

$$\int [\mathrm{Im}\, c_n(t)]^2\, dG(t)$$

for testing symmetry of F. When the center of symmetry is not specified, a modified statistic such as

$$\inf_\mu \int [\mathrm{Im}\, \{e^{i\mu t} c_n(t)\}]^2\, dG(t)$$

may be considered.

Received January 1974; revised March 1976.

[1] Research supported by National Research Council of Canada Grant A-9133.

AMS 1970 *subject classifications.* Primary 62G99, 60G99, 62M99.

Key words and phrases. Empirical characteristic function, characteristic function, uniform almost sure convergence, weak convergence, Gaussian processes, testing for symmetry, asymptotic distribution.

The range of problems to which the ecf seems applicable appears to be quite wide. This is because Fourier–Stieltjes transformation often results in easy translation of properties that are important in problems of inference. The characteristic function behaves simply under shifts, scale changes and summation of independent variables; it allows an easy characterization of independence and of symmetry. It is therefore not difficult to suggest ecf procedures for areas of inference such as testing for goodness of fit, testing for independence, parameter estimation, etc. Further, although it is true that there exist other functionals in one-to-one correspondence with the distributions, it may be argued that the characteristic function is a uniquely important concept for which a large body of theory is available. The ecf retains all information present in the sample and lends itself conveniently to computation. Finally, there are situations where a characterization of some property or of a class of distributions exists in terms of characteristic functions. One example is the problem of inference on the parameters of the stable laws. Here traditional methods have not led to a solution and an ecf approach seems likely to lead to useful procedures. See for example Paulson, Halcomb and Leitch (1975).

A search of the literature reveals some scattered references to the ecf. Its definition is given by Parzen (1962) and it is used for statistical inference by Heathcote (1972) and Press (1972). Work on statistical applications appears to be impeded because not enough is known about the ecf.

2. Convergence of the ecf. For any fixed t, $c_n(t)$ is an average of bounded i.i.d. random variables having means $c(t)$. Therefore it follows by the strong law of large numbers that $c_n(t)$ converges almost surley to $c(t)$. Further:

THEOREM 2.1. *For fixed* $T < \infty$,

$$P\{\lim_{n\to\infty} \sup_{|t|\leq T} |c_n(t) - c(t)| = 0\} = 1 .$$

PROOF. According to the Glivenko–Cantelli theorem (Loève (1963), page 20) it follows with probability one that F_n converges completely (ibid., page 178) to F. The result follows upon applying the criteria (ibid., page 191) of complete convergence and uniform convergence. □

Because $c_n(t)$ is a trigonometric polynomial, it is almost periodic and hence must approach its supremum value $c_n(0) = 1$ arbitrarily often as $|t| \to \infty$. (See, for example Bohr (1947), pages 38, 32, 31.) On the other hand, we may have $c(t) \to 0$ as $|t| \to \infty$, as for example when $F(x)$ is absolutely continuous (see Lukacs (1970), page 19). Therefore the conclusion of Theorem 2.1 cannot generally be true for $T = \infty$. If however $F(x)$ where supported on a lattice $0, \pm a, \pm 2a, \cdots$, then both $c(t)$ and $c_n(t)$ will have period $2\pi a^{-1}$ and the result would hold with $T = \infty$. More generally:

THEOREM 2.2. *Let* $F(x)$ *be purely discrete. Then*

$$P\{\lim_{n\to\infty} \sup_{-\infty<t<\infty} |c_n(t) - c(t)| = 0\} = 1 .$$

PROOF. This follows from the strong law of large numbers. If X is a discrete random variable taking values $\{a_k\}$ with probabilities $\{p_k\}$ then in obvious notation

$$\begin{aligned}|c_n(t) - c(t)| &= |\sum_k (p_{k,n} - p_k)e^{ita_k}| \\ &\leqq \sum_k |p_{k,n} - p_k| \to 0 \quad \text{a.s.} \end{aligned} \qquad \square$$

The connection (see Parzen (1962)) between the ecf and kernel type estimators of probability density functions (see for example Wegman (1972)) makes it possible to apply results in density estimation to the ecf. Consider for example the following result of Nadarya (1965):

THEOREM 2.3. *Let $F(x)$ be absolutely continuous with uniformly continuous derivative $f(x)$. Let $K(x)$ be a probability density of bounded variation and $\{a(n)\}$ be a sequence of nonnegative numbers converging to zero and satisfying*

$$\sum_{n=1}^{\infty} e^{-\gamma n a^2(n)} < \infty \qquad \textit{for each} \quad \gamma > 0 .$$

Let $f_n(x) = (1/na(n)) \sum_{j=1}^{n} K((x - X_j)/a(n))$. Then

$$P\{\lim_{n\to\infty} \sup_{-\infty<x<\infty} |f_n(x) - f(x)| = 0\} = 1 .$$

As a consequence of Nadarya's result we can prove:

THEOREM 2.4. *Let $F(x)$ have characteristic function $c(t) \to 0$ as $|t| \to \infty$. Let $K(x)$, $a(n), f_n(x)$ be as in Theorem 2.3, let $\hat{K}$ be the Fourier transform of K, and denote by $\hat{c}_n(t) = \hat{K}(a(n)t)c_n(t)$ the characteristic function of $f_n(x)$. Then*

$$P\{\lim_{n\to\infty} \sup_{-\infty<x<\infty} |\hat{c}_n(t) - c(t)| = 0\} = 1 .$$

We need the following lemma which may be proved by standard arguments:

LEMMA 2.5. *Let W_n be a sequence of random variables. Then $W_n \to 0$ a.s. if and only if there exists a sequence $0 < \delta_j \to 0$ such that*

$$P\{\bigcap_{j=n}^{\infty} [|W_j| \leqq \delta_j]\} \to 1 \qquad \textit{as} \quad n \to \infty .$$

PROOF OF THEOREM 2.4. First suppose F is absolutely continuous with uniformly continuous derivative $f(x)$. Now

$$\sup_t |c_n(t) - c(t)| \leqq A_n + B_n + C_n$$

where

$$A_n = \int_{-M_n}^{M_n} |f_n'(x) - f(x)|\, dx ,$$

$$B_n = \int_{|x|>M_n} |f(x)|\, dx , \qquad C_n = \int_{|x|>M_n} |f_n(x)|\, dx .$$

To choose M_n, apply Lemma 2.5 to Theorem 2.3 with $W_j = \sup_x |f_j(x) - f(x)|$ to obtain a sequence $0 < \delta_j \to 0$ and set $M_n = \delta^{-\frac{1}{2}}$. Then $B_n \to 0$; $C_n \to 0$ a.s. (by the strong law of large numbers); and finally $A_n \to 0$ a.s. by Lemma 2.5 and

$$P\{\bigcap_{j=n}^{\infty} [|A_j| \leqq 2M_j\delta_j]\} \geqq P\{\bigcap_{j=n}^{\infty} [|W_j| \leqq \delta_j]\} \to 1 .$$

The restriction on F may be removed by a convolution argument: Replace X_j by $X_j^* = X_j + U_j$, $j = 1, 2, \cdots, n$ where the U_j are i.i.d. with density function

$$\begin{array}{ll} \delta^{-1}(1-\delta^{-1}|x|)\,, & |x| \leqq \delta \\ 0\,, & \text{otherwise} \end{array}$$

and characteristic function $\psi_\delta(t)$. Let c^*, c_n^* be the ch.f., ecf corresponding to the X_j^* and choose $S(\varepsilon)$ so that $|K(\hat{t})| \leqq \varepsilon$ when $|t| > S(\varepsilon)$. The result follows from the inequality

$$\begin{aligned} \sup_t |\hat{c}_n(t) - c(t)| \leqq{} & \sup_t |K(a(n)t)(c_n(t) - c_n^*(t))| \\ & + \sup_t |\hat{K}(a(n)t)c_n^*(t) - c^*(t)| + \sup_t |c^*(t) - c(t)|\,. \end{aligned}$$

The last term may be bounded, say by ε, by taking δ sufficiently small because $c^*(t) = \psi_\delta(t)c(t)$, $c(t) \to 0$ as $|t| \to \infty$, and $\psi_\delta(t) \to 1$ as $\delta \to 0$ uniformly in bounded intervals. The middle term converges a.s. for fixed $\delta > 0$ because the convolution X_j^* has uniformly continuous density. Finally the first term is bounded by

$$\begin{aligned} & \sup_{|t| \geqq S(\varepsilon)} |\hat{K}(a(n)t)(c_n(t) - c_n^*(t))| + \sup_{|t| < S(\varepsilon)} \left|\frac{1}{n} \textstyle\sum_{j=1}^n (e^{itX_j} - e^{it(X_j + U_j)})\right| \\ & \qquad \leqq 2\varepsilon + \frac{1}{n} S(\varepsilon) \textstyle\sum_{j=1}^n |U_j| \leqq 3\varepsilon \end{aligned}$$

provided that $\delta < \varepsilon/S(\varepsilon)$. □

We have also:

THEOREM 2.6. *Let F be an arbitrary distribution function whose singular part has characteristic function vanishing at the extremities. Let $T_n = 0((n/\log n)^{\frac{1}{2}})$. Then*

$$P\{\lim_{n\to\infty} \sup_{|t| \leqq T_n} |c_n(t) - c(t)| = 0\} = 1\,.$$

PROOF. Assume first that F has no discrete part so that it satisfies the condition of Theorem 2.4.

According to a result of Mureika (1972), if $\{u(n)\}$ is a sequence of nonnegative numbers then

$$\begin{aligned} \textstyle\sum_{n=1}^\infty e^{-u(n)x} &< \infty\,, && \text{if } x > \lambda \\ &= +\infty\,, && \text{if } x < \lambda \end{aligned}$$

where $\lambda = \limsup\,(u^{-1}(n) \log n)$. Choose $\{a(n)\}$ so that $a(n)(n/\log n)^{\frac{1}{2}} \to \infty$ and $a(n)T_n \to 0$. It follows that this sequence will satisfy the condition of Theorem 2.3.

Set $\hat{c}_n(t) = K(a(n)t)c_n(t)$ as before, and consider the inequality

$$\sup_{|t| \leqq T_n} |c_n(t) - c(t)| \leqq \sup_{|t| \leqq T_n} |c_n(t) - \hat{c}_n(t)| + \sup_{|t| \leqq T_n} |\hat{c}_n(t) - c(t)|\,.$$

The last term on the right converges according to Theorem 2.4. The first is bounded by

$$\sup_{|t| \leqq T_n} |1 - \hat{K}(a(n)t)|$$

and $\hat{K}$ is continuous, $K(0) = 1$, and $|a(n)t| \leqq a(n)T_n \to 0$.

When F has a discrete component, $c_n(t) - c(t)$ may be decomposed in a natural way into its discrete and continuous parts. The result follows on applying Theorem 2.2 to the discrete part. □

The following theorem gives another type of convergence result for the ecf:

THEOREM 2.7. *Let F be any distribution function. Let $T_n = 0(n^{p/2})$ where $0 < p \leqq 2$. Then*

$$V_n = \int_0^{T_n} |c_n(t) - c(t)|^p \, dt \to 0 \quad \textit{in probability.}$$

PROOF. For nonnegative random variables V we have the inequalities $P\{V > \varepsilon\} \leqq \varepsilon^{-1}EV$ and $EV^p \leqq (EV^2)^{p/2}$, $0 < p \leqq 2$. (See for example Loève (1963), Section 9.3.) Then

$$EV_n \leqq \int_0^{T_n} \{E|c_n(t) - c(t)|^2\}^{p/2} \leqq n^{-p/2}T_n \to 0$$

and the result follows. □

3. The ecf as a stochastic process. Let $c_n(t) = (1/n) \sum_{j=1}^n e^{itx_j}$ and consider $Y_n(t) = n^{\frac{1}{2}}(c_n(t) - c(t))$ as a random complex process in t. It may be seen that $EY_n(t) = 0$ and $EY_n(t_1)Y_n(t_2) = c(t_1 + t_2) - c(t_1)c(t_2)$, this latter term fully determining the covariance structure of $Y_n(t)$. Define $Y(t)$ to be a zero mean complex valued Gaussian process satisfying $\bar{Y}(t) = Y(-t)$ and having the same covariance structure as $Y_n(t)$. Note that

$$\operatorname{Cov}(\operatorname{Re} Y(t_1), \operatorname{Re} Y(t_2)) = \tfrac{1}{2}[\operatorname{Re} c(t_1 + t_2) + \operatorname{Re} c(t_1 - t_2)] - \operatorname{Re} c(t_1) \operatorname{Re} c(t_2) ,$$
$$\operatorname{Cov}(\operatorname{Re} Y(t_1), \operatorname{Im} Y(t_2)) = \tfrac{1}{2}[\operatorname{Im} c(t_1 + t_2) + \operatorname{Im} c(t_1 - t_1)] - \operatorname{Re} c(t_1) \operatorname{Im} c(t_2) ,$$
$$\operatorname{Cov}[\operatorname{Im} Y(t_1), \operatorname{Im} Y(t_2)) = \tfrac{1}{2}[-\operatorname{Re} c(t_1 + t_2) + \operatorname{Re} (t_1 - t_2)] - \operatorname{Im} c(t_1) \operatorname{Im} c(t_2) .$$

For finite collections $t_1, t_2, \cdots, t_m$, application of the multidimensional central limit theorem implies convergence in distribution of $Y_n(t_1), Y_n(t_2), \cdots, Y_n(t_m)$ to $Y(t_1), Y(t_2), \cdots, Y(t_m)$. More generally:

THEOREM 3.1. *Let $Y_n(t)$, $Y(t)$ be as defined above. The process $Y_n(t)$ converges weakly to $Y(t)$ in every finite interval.*

PROOF. According to Corollary 7 of Whitt (1970), measures P_n, $n \geqq 1$ and P on $C^2[0, 1]$ satisfy the weak convergence $P_n \Rightarrow P$ if and only if (i) the finite dimensional distributions of P_n converge weakly to the finite dimensional distributions of P and (ii) the two marginal measures on $C[0, 1]$ are tight.

Now,

$$\begin{aligned} E|Y_n(t_2) - Y_n(t_1)|^2 &\leqq 2[1 - \operatorname{Re} c(t_1 - t_2)] \\ &= 2 \int_{-\infty}^{\infty} [1 - \cos x(t_2 - t_1)] \, dF(x) \\ &\leqq 2 \int_{-\infty}^{\infty} |x(t_2 - t_1)|^{1+\delta} \, dF(x) \\ &= 2|t_2 - t_1|^{1+\delta} E|X|^{1+\delta} , \qquad 0 \leqq \delta \leqq 1 . \end{aligned}$$

Hence tightness of the real and imaginary parts follow from Theorem 12.3 of Billingsley (1968) in the case that $E|X|^{1+\delta}$. The moment condition is removed by a truncation type argument using Theorem 4.2 of Billingsley (1968). □

An interesting and related weak convergence result for the ecf is proved by Kent (1975).

4. Application to testing for symmetry. In this section an application of the ecf to the problem of testing for symmetry about the origin is outlined. Although the ecf does lead to reasonable procedures here, no attempt is made at comparison with other results in the area of symmetry testing.

The statistic proposed is

$$\begin{aligned} T_n &= \int_{-\infty}^{\infty} [\operatorname{Im} c_n(t)]^2 \, dG(t) \\ &= 2^{-1} n^{-1} \sum_{j=1}^{n} \sum_{k=1}^{n} [g(X_j - X_k) - g(X_j + X_k)] \,. \end{aligned}$$

Here G is taken to be a distribution function symmetric about the origin and having characteristic function $g(u)$. Concerning the distribution of T_n we have:

THEOREM 4.1. *Let $I = \int [\operatorname{Im} c(t)]^2 \, dG(t)$ and*

$$\begin{aligned} \sigma^2 = \iint \operatorname{Im} c(t_1) \operatorname{Im} c(t_2) \{2 \operatorname{Re} [-c(t_1 + t_2) + c(t_1 - t_2)] \\ - 4 \operatorname{Im} c(t_1) \operatorname{Im} c(t_2)\} \, dG(t_1) \, dG(t_2) \,. \end{aligned}$$

Then when $\sigma^2 > 0$, $n^{\frac{1}{2}}(T_n - I)$ is asymptotically normal with mean 0 and variance σ^2. When $\sigma^2 = 0$ the asymptotic distribution is degenerate at 0.

PROOF. Consider the equality

$$n^{\frac{1}{2}}(T_n - I) = \tfrac{1}{2} n^{-\frac{1}{2}} \sum_{j=1}^{n} [1 - g(2X_j)] + \tfrac{1}{2} \frac{n-1}{n} n^{\frac{1}{2}}(U_n - 2I) - n^{-\frac{1}{2}} I$$

where $U_n = \binom{n}{2}^{-1} \sum_c [g(X_j - X_k) - g(X_j + X_k)]$, this sum being over the $\binom{n}{2}$ combinations. The first and third terms on the right converge to 0. Regarding the middle term, U_n is a Hoeffding U-statistic and is asymptotically normal with mean

$$E[g(X_1 - X_2) - g(X_1 + X_2)] = 2I$$

and variance

$$n^{-1} \operatorname{Var} \{E_{X_2}[g(X_1 - X_2) - g(X_1 + X_2)]\} = n^{-1}\sigma^2 \,.$$

(See Fraser (1957), Section 6.5 for example.) □

The inequalities $0 \leqq I \leqq 1$ and $0 \leqq \sigma^2 \leqq 4$ may be noted. Under symmetry $I = \sigma^2 = 0$. Because there exist nonsymmetric characteristic functions that are purely real in a finite neighbourhood of the origin, it is possible to have $\sigma^2 = 0$ for a nonsymmetric F.

Concerning the distribution of T_n under the null hypothesis we have:

THEOREM 4.2 *Let F be symmetric about the origin. Let $W_n(t) = n^{\frac{1}{2}} \sum_{j=1}^{n} \sin tX_j$ and let $W(t)$ be a zero mean Gaussian process having the same covariance function as $W_n(t)$:*

$$K(t_1, t_2) = \tfrac{1}{2}[-c(t_1 + t_2) + c(t_1 - t_2)] \,.$$

Then $nT_n = \int_{-\infty}^{\infty} W_n^2(t)\,dG(t)$ *converges in distribution to* $\xi = \int_{-\infty}^{\infty} W^2(t)\,dG(t)$.

PROOF. This result follows on applying Theorem 4.2 of Billingsley (1968) to

$$X_u = \int_{-u}^{u} W^2(t)\,dG(t)\,, \qquad X = \int_{-\infty}^{\infty} W^2(t)\,dG(t)$$
$$X_{un} = \int_{-u}^{u} W_n^2(t)\,dG(t)\,, \qquad Y_n = \int_{-\infty}^{\infty} W_n^2(t)\,dG(t)\,.$$

The lim lim sup condition follows from the Markov inequality (Loève (1963), page 158):

$$P\{\int_{|t|>u} W_n^2(t)\,dG(t) \geqq \varepsilon\} \leqq \varepsilon^{-2}\int_{|t_1|>u}\int_{|t_2|>u} EW_n^2(t_1)W_n^2(t_2)\,dG(t_1)\,dG(t_2)\,.$$

A similar argument gives $X_u \to_P X$. □

The following result gives the distribution of ξ in the case that $G(x)$ is absolutely continuous with density function $G'(x) = g_*(x)$ such that $g_*(x) = 0$ for $|x| > M < \infty$, and $g_*(x)$ is continuous on $[-M, M]$:

THEOREM 4.3. *The characteristic function of the random variable*

$$\xi = \int_{-M}^{M} W^2(t)g_*(t)\,dt$$

is given by

$$\phi(t) = \prod_{j=1}^{\infty}(1 - 2it\lambda_j)^{-\frac{1}{2}}$$

where $\{\lambda_j\}$ *is the solution set of the eigenvalue equation*

$$\lambda_j\psi_j(t) = \int_{-M}^{M} \psi_j(s)K(s, t)(g_*(s)g_*(t))^{\frac{1}{2}}\,ds\,.$$

PROOF. According to the theorem of Karhunen–Loève (see the appendix in Ash (1965), for example)

$$W(t) = \sum_{j=1}^{\infty} Z_j\psi_j(t)\,, \qquad -M \leqq t \leqq M\,,$$

the convergence being in mean square and uniformly in t, and

$$Z_j = \int_{-M}^{M} W(t)\psi_j(t)\,dt$$

being independent normal variables having means 0 and variances λ_j. The ψ_j are taken orthonormal, and $\sum \lambda_j < \infty$. It follows that

$$\xi = \sum_{j=1}^{\infty} Z_j^2$$

is distributed as a sum of independent $\lambda_j\chi_1^2$ variables. □

The restriction that G have bounded support is less troublesome than might first be thought. There are two reasons for this. Firstly, in practice, the realized X_j values are of necessity discretized with recorded values typically confined to a grid $0, \pm\Delta\ \pm 2\Delta, \cdots$ (Δ very small); the grid point chosen in each case presumably being that closest to the "actual" value. Hence the "actual" characteristic function is neither estimable nor relevant for frequencies of order of magnitude greater than Δ^{-1}. This may be compared to the phenomena of sampling and aliasing in time series analysis. Secondly, as a consequence of analyticity (Theorem 7.2.1. of Lukacs (1970) for example) the characteristic function of a bounded random variable is determined by its values in any finite interval containing 0. And the problem may always be reduced to one of bounded random variables: Let $H(x)$ be any absolutely continuous, strictly increasing distribution function which is symmetric about the origin. Then X is symmetric about the origin if and only if $2H(X) - 1$ is.

The distribution of a weighted sum of independent chi-squared variates may be found by numerical integration of the characteristic function with the aid of the fast Fourier transform. Alternatively, a well-known approximation is to use a multiple $\theta\chi_\nu^2$ of a chi-squared with θ, ν chosen to match the first two moments. These may be computed directly from $\xi = \int_{-M}^{M} W^2(t)\, dG(t)$:

$$E\xi = \tfrac{1}{2} \int_{-M}^{M} [1 - c(2t)]\, dG(t)$$

and

$$\begin{aligned}\operatorname{Var}(\xi) &= \int_{-M}^{M} \int_{-M}^{M} \operatorname{Cov}(W^2(t_1), W^2(t_2))\, dG(t_1)\, dG(t_2) \\ &= 2 \int_{-M}^{M} \int_{-M}^{M} K^2(t_1, t_2)\, dG(t_1)\, dG(t_2)\,.\end{aligned}$$

In general c will be unknown, but may be estimated by

$$c_R(t) = \frac{1}{n} \sum_{j=1}^{n} \cos tX_j\,.$$

When the underlying distribution is symmetric, $c_R(t)$ will become uniformly close to $c(t)$. This leads to a test procedure which, to within the accuracy of the $\theta\chi_\nu^2$ approximation, will have asymptotic level α.

An alternative approach might be to estimate the $\{\lambda_j\}$. Using the estimate $c_R(t)$ the eigenvalue equation becomes

$$\lambda_j \phi_j(t) = \frac{1}{n} \sum_{k=1}^{n} \gamma_{kj} (g_*(t))^{\frac{1}{2}} \sin(tX_k)$$

where

$$\gamma_{kj} = \int_{-M}^{M} (g_*(s))^{\frac{1}{2}} \sin(sX_k) \phi_j(s)\, ds\,,$$

or, on substituting the first equation into the second,

$$\lambda_j \gamma_{kj} = \frac{1}{n} \sum_{l=1}^{n} \gamma_{lj} D_{kl}$$

where

$$D_{kl} = \tfrac{1}{2}[g(X_k - X_l) - g(X_k + X_l)]\,,$$

g being the characteristic function of G. This is now an $n \times n$ eigenvalue problem. Once the eigenvalues $\lambda_1^{(n)}, \lambda_2^{(n)}, \cdots, \lambda_n^{(n)}$ of $\{D_{kl}\}$ are determined, the characteristic function of ξ is estimated as

$$\phi_n(t) = \prod_{j=1}^{n} (1 - 2it\lambda_j^{(n)})^{-\frac{1}{2}}\,.$$

An alternate form in terms of the characteristic polynomial of the matrix is:

$$\phi_n(t) = (-2it)^{-n/2} \left(\det\left[\{D_{kl}\} - \frac{1}{2it}\mathbf{I}\right]\right)^{-\frac{1}{2}}\,.$$

$\phi_n(t)$ may then be inverted to approximate the distribution of ξ as a finite linear combination of χ_1^2 variates. This procedure is justified by the following theorem and leads to a test procedure with asymptotic level α:

THEOREM 4.4. *Let E_n, E be the cumulative distribution functions corresponding to ϕ_n, ϕ. Then with probability one*

$$\sup_x |E_n(x) - E(x)| \to 0 .$$

PROOF. To simplify notation, write $K(s, t)$ for $K(s, t)(g_*(s)g_*(t))^{\frac{1}{2}}$, and similarly write $K_n(s, t)$ for the kernel based on $c_R(t)$. The following assertions each hold almost surely: $c_R(t)$ converges uniformly to $c(t)$ in every bounded interval. Hence $K_n(s, t)$ converges uniformly to $K(s, t)$ in every bounded region. By the property (see page 151 of Courant and Hilbert (1953) for example) of continuous dependence on the kernel,

$$\lambda_j^{(n)} \to \lambda_j \qquad \text{as} \quad n \to \infty\,, \quad \text{all} \quad j$$

where all eigenvalue sets are assumed in decreasing order. This property, together with the positiveness and uniform boundedness of the quantities $\sum \lambda_j^{(n)} = \int K_n(t, t)\, dt$ and $\sum \lambda_j = \int K(t, t)\, dt$ suffice to establish, after a little analysis, the convergence

$$\prod_{j=1}^{n} (1 - 2it\lambda_j^{(n)})^{-\frac{1}{2}} \to \prod_{j=1}^{\infty} (1 - 2it\lambda_j)^{-\frac{1}{2}}$$

uniformly on $(-\infty, \infty)$. The result now follows from the complete convergence criterion of page 191, Loève (1963). □

The authors conjecture that Theorems 4.3 and 4.4 remain valid for $M = \infty$ and that the continuity requirement on $g_*(t)$ may be relaxed.

Acknowledgment. The authors would like to thank the editor for many suggestions which helped very much to improve this paper.

REFERENCES

ASH, R. B. (1965). *Information Theory*. Wiley-Interscience, New York.

BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York.

BOHR, H. A. (1947). *Almost Periodic Functions*. Chelsea, New York.

COURANT, R. and HILBERT, D. (1953). *Methods of Mathematical Physics* **1**. Interscience, New York.

FRASER, D. A. S. (1957). *Nonparametric Methods in Statistics*. Wiley, New York.

HEATHCOTE, C. R. (1972). A test of goodness of fit for symmetric random variables. *Austral. J. Statist.* **14** 172–181.

KENT, J. T. (1975). A weak covergence theorem for the empirical characteristic function. *J. Appl. Probability* **12** 515–523.

LOÈVE, M. (1963). *Probability Theory*. Van Nostrand, New York.

LUKACS, E. (1970). *Characteristic Functions*. Griffin, London.

MUREIKA, R. A. (1972). The maximization of entropy of discrete denumerably valued random variables with known mean. *Ann. Math. Statist.* **43** 541–552.

NADARYA, E. A. (1965). On non-parametric estimates of density functions and regression curves. *Theor. Probability Appl.* **10** 186–190.

Parzen, E. (1962). On estimation of a probability density function and mode. *Ann. Math. Statist.* **33** 1065–1076.

PAULSON, A. S., HALCOMB, E. W. and LEITCH, R. A. (1975). The estimation of the parameters of the stable laws. *Biometrika* **62** 163–170.

PRESS, S. J. (1972). Estimation in univariate and multivariate stable distributions. *J. Amer. Statist. Assoc.* **67** 842–846.

WEGMAN, E. J. (1972). Nonparametric probability density estimation: I. A summary of available methods. *Technometrics* **14** 533–546.

WHITT, W. (1970). Weak convergence of probability measures on the function space $c[0, \infty)$. *Ann. Math. Statist.* **41** 939–944.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF TORONTO
TORONTO, CANADA M5S 1A1

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF ALBERTA
EDMONTON, ALBERTA T6G 2E2