The Annals of Mathematical Statistics
1969, Vol. 40, No. 3, 854–864

EFFICIENT ESTIMATION OF A PROBABILITY DENSITY FUNCTION

BY JAMES PICKANDS III

Virginia Polytechnic Institute

1. Introduction. The problem of estimating a probability density function has received a considerable amount of attention in recent years. In particular, there has been extensive interest in estimators based on weight functions. Specifically, let

$$\hat{f}_n(x) \equiv n^{-1} \sum_{j=1}^{n} \Psi_n(x - X_j), \tag{1.1}$$

where for each n, $\Psi_n(x)$ is a square integrable function, and X_j, $1 \leqq j \leqq n$, are the sample values. See, for example, Parzen [6], and Watson and Leadbetter [7]. To evaluate an estimator a "figure of merit" is needed. The usual one is I_n^2, where

$$I_n^2 \equiv 2\pi E \int_{-\infty}^{\infty} |\hat{f}_n(x) - f(x)|^2 \, dx.$$

If $f(x) \varepsilon L_2$, then I_n^2 is necessarily finite, since $\hat{f}_n(x) \varepsilon L_2$, by the assumption that the functions $\Psi_n(x) \varepsilon L_2$. By the Parseval Identity,

$$I_n^2 = E \int_{-\infty}^{\infty} |\hat{\phi}_n(t) - \phi(t)|^2 \, dt,$$

where $\hat{\phi}_n(t)$ and $\phi(t)$ are the Fourier transforms of $\hat{f}_n(x)$ and $f(x)$ respectively. For this reason in studying the asymptotic efficiency, it is convenient to work with Fourier transforms. But

$$\hat{\phi}_n(t) = A_n(t)\tilde{\phi}_n(t)$$

where

$$\tilde{\phi}_n(t) \equiv n^{-1} \sum_{j=1}^{n} e^{iX_j t}.$$

and $A_n(t)$ is the Fourier transform of $\Psi_n(x)$. We call this the empirical characteristic function. Since $\hat{\phi}_n(t)$ is not necessarily positive definite, $\hat{f}_n(x)$ is not necessarily non-negative. Observe, though, that I_n^2 can be reduced by replacing $\hat{f}_n(x)$ with 0, whenever it is negative.

If $\lim_{n\to\infty} I_n^2 = 0$, we say that the estimator sequence is consistent, or more explicitly, that it is consistent in mean square. Conditions for consistency are well known for those estimators which are commonly used. See [7].

For any given characteristic function $\phi(t)$, the function $A_n(t)$ can be so chosen as to minimize I_n^2, as shown in [7]. We define

$$J_n^2 \equiv \min I_n^2,$$

where the minimum is taken over all possible functions $A_n(t)$. This makes possible

Received 28 March 1968.