NONPARAMETRIC ESTIMATION OF A DENSITY FUNCTIONAL

by

Ibrahim A. Ahmad, Memorial University of Newfoundland

and

Pi-Erh Lin, Florida State University

FSU Statistics Report M369

The Florida State University
Tallahassee, Florida 32306
January, 1976
NONPARAMETRIC ESTIMATION OF A DENSITY FUNCTIONAL

by

Ibrahim A. Ahmad, Memorial University of Newfoundland

and

Pi-Erh Lin, Florida State University

ABSTRACT

Let \( X \) be a random variable with distribution function \( F \) and density function \( f \). Let \( \phi \) and \( \psi \) be known measurable functions defined on the real line \( \mathbb{R} \) and the closed interval \( [0, 1] \), respectively. This paper proposes a smooth nonparametric estimate of the density functional \( \theta = \int \phi(x)\psi(F(x))f^2(x)dx \) based on a random sample \( X_1, \ldots, X_n \) from \( F \) using a kernel function \( k \). The proposed estimate is given by

\[
\hat{\theta} = (n^2a_n)^{-1} \sum_{i=1}^{n} \sum_{j=1}^{n} \phi(X_i)\psi(\hat{F}(X_i))k[(X_i - X_j)/a_n],
\]

where \( \{a_n\} \) is a sequence of positive real numbers converging to 0, as \( n \to \infty \), and where \( \hat{F}(x) = n^{-1}\sum_{i=1}^{n} K[(x - X_i)/a_n] \) with \( K(w) = \int_{-\infty}^{w} k(u)du \). The estimate \( \hat{\theta} \) is shown to be consistent both in the weak and strong senses. Conditions are obtained under which the asymptotic normality of \( n^{1/2}(\hat{\theta} - \theta) \) is established. Applications of the estimate to the study of asymptotic relative efficiency for various nonparametric tests are indicated, with particular reference to those using the Chernoff-Savage statistic.


Key words and phrases: Kernel function; strong consistency; asymptotic normality; asymptotic relative efficiency; Chernoff-Savage statistic.
NONPARAMETRIC ESTIMATION OF A DENSITY FUNCTIONAL

by

Ibrahim A. Ahmad, Memorial University of Newfoundland

and

Pi-Erh Lin¹, Florida State University

1. Introduction. Let $X$ be a random variable with distribution function (df) $F$ and probability density function (pdf) $f$. Let $\phi$ and $\psi$ be known measurable functions defined on the real line $R = (-\infty, \infty)$ and the closed interval $[0, 1]$, respectively. Consider the density functional

\begin{equation}
\theta = \theta(\phi, \psi, f, F) = \int \phi(x) \psi[F(x)] f^2(x) dx.
\end{equation}

(Here and throughout the study no limits of integration are given whenever the integration extends from $-\infty$ to $\infty$.) The functional $\theta$ is useful in many applications. For example, special cases of $\theta$ appear as dominant terms in the asymptotic relative efficiency (ARE) of many nonparametric tests. These ARE's are unknown quantities when nothing is known about $F$. Thus a nonparametric estimate of $\theta$ is needed to provide information about the relative performance of two rival tests under suitable sequences of alternative hypotheses.

¹Research supported in part by the National Institute of General Medical Sciences through the Training Grant 5T01 GM-913.
with \( K(w) = \int_{-\infty}^{\infty} k(z)dz \), and

\[
\hat{f}(x) = \frac{1}{n} \int_{-\infty}^{\infty} k\left(\frac{x-y}{a_n}\right) dF_n(y) = \frac{1}{na_n} \sum_{j=1}^{n} k\left(\frac{x-x_j}{a_n}\right).
\]

The expressions of \( \hat{F}(x) \) and \( \hat{f}(x) \) given above are known as the kernel estimates of df and pdf, respectively.

Sen (1966), in studying confidence intervals for a shift parameter and the ARE's of nonparametric tests, has proposed a general method of estimating \( \theta(1, J', f, F) \) and \( \theta(x, J', f, F) \) using rank order statistics. (Here \( J' \) is the derivative of \( J \)). Among other results, he has obtained the asymptotic normality and weak consistency of his estimates. In a different context, Bhattacharaya and Roussas (1969) have proposed to estimate \( \theta(1, 1, f, F) \) by \( \int \hat{f}^2(x)dx \) and showed that their estimate is consistent in the first and second means. Also using a kernel function, Schuster (1971) proposes to estimate \( \theta(1, 1, f, F) \) by \( \int \hat{f}(x)dF_n(x) \) and obtains the convergence rate for his estimate. However, no limiting distribution is established for either estimate of \( \theta(1, 1, f, F) \).

In Section 2 it is shown that \( \hat{\theta} \) converges to \( \theta \) in both the weak and strong senses. Asymptotic normality of \( n^{1/2}(\hat{\theta} - \theta) \) is established in Section 3. Section 4 gives some remarks on the functional \( \theta \), with particular reference to the ARE's of nonparametric tests using the Chernoff-Savage statistic.
Then

\[|\hat{\theta} - \theta| \leq I_{1n} + I_{2n} + I_{3n} + I_{4n}.\]

**THEOREM 2.1.** Assume the following conditions

(i) \(k\) is uniformly continuous satisfying (1.2);
(ii) \(n_{2n} \to \infty, \text{ as } n \to \infty;\)
(iii) \(C = \sup_{0 \leq t \leq 1} \left| \frac{d}{dt} \psi(t) \right| < \infty;\)
(iv) \(E[\psi(X)\psi(F(X))f(X)]^2 < \infty;\)
(v) \(E[\sup_x \hat{f}(x) - Ef(x)]^2 \to 0, \text{ as } n \to \infty;\)
(vi) \(E(G^4) \to 0 \text{ as } n \to \infty, \text{ where } G = \sup_x |\hat{F}(x) - F(x)|;\)
(vii) \(E\Phi^2(X) < \infty;\)
(viii) \(E[\psi(X)\psi(F(X))]^2 < \infty.\)

Then

\[E|\hat{\theta} - \theta| \to 0, \text{ as } n \to \infty.\]

**PROOF.** It suffices to show that, for \(\alpha = 1, 2, 3, 4,\)

\[E I_{\alpha n} \to 0, \text{ as } n \to \infty.\]

Since \(I_{4n}\) is a constant and since
where

\[(2.11) \quad D = \mathbb{E}\{\mathbb{E}[\text{sup}_{x} \hat{f}(X_1)|X_1 = x]\}^{1/4} \}.
\]

It follows from Condition (v) that \(\text{sup} \hat{f}(x)\) is bounded in probability. This implies that \(D < \infty\) for all \(n\). Thus, by Conditions (vi) and (vii), \(E_{1n} \to 0\) as \(n \to \infty\).

For \(\alpha = 2\), we have

\[(2.12) \quad E_{2n} \leq \mathbb{E}\left|\phi(x)\psi[F(x)]\right| \mathbb{E}[\hat{f}(x) - \mathbb{E}\hat{f}(x)]dF_n(x) \]

\[\leq \mathbb{E}[\text{sup}_{z} \hat{f}(z) - \mathbb{E}\hat{f}(z)]^2 \mathbb{E}\left[\mathbb{E}[\phi(x)\psi[F(x)]]dF_n(x)\right]^2.\]

In view of (v) we need only show that the second factor converges to a finite quantity. Now

\[(2.13) \quad \mathbb{E}\{\left|\phi(x)\psi[F(x)]\right| dF_n(x)\}^2 \]

\[= \mathbb{E}\left\{\frac{1}{n} \sum_{i=1}^{n} \phi(X_i)\psi[F(X_i)]\right\}^2 \]

\[= \text{Var} \{ \} + \mathbb{E}\{ \}^2 \]

\[= \frac{1}{n} \mathbb{E}\{\phi(X_1)\psi[F(X_1)]\}^2 + \frac{n-1}{n} \mathbb{E}\{\phi(X_1)\psi[F(X_1)]\}^2 \]

\[\to \{\mathbb{E}\phi(X_1)\psi[F(X_1)]\}^2 < \infty, \]
THEOREM 2.2. Assume (i) - (iv) of Theorem 2.1 and the following conditions:

(i) \( G = \sup_{x} |\hat{F}(x) - F(x)| \to 0 \), w.p.1, as \( n \to \infty \);

(ii) \( E[|\phi(X)|f(X)] < \infty \);

(iii) \( \sup_{x} |\hat{F}(x) - f(x)| \to 0 \), w.p.1, as \( n \to \infty \);

(iv) \( a_{n}^{-1} \sup_{x} \left| F_{n}(x) - F(x) \right| \to 0 \), w.p.1, as \( n \to \infty \);

(v) \( \mu = \int |\phi(x)|f(x) < \infty \); and

(vi) \( E[|\phi(X)|F(X)] E[\hat{f}(X)|X] < \infty \).

Then

\[ |\hat{\theta} - \theta| \to 0 \), w.p.1, as \( n \to \infty \).

PROOF. Since \( I_{\alpha n} \to 0 \) as \( n \to \infty \), it suffices to show that, for \( \alpha = 1, 2, \) and \( 3 \),

\[ (2.14) \quad I_{\alpha n} \to 0 \), w.p.1, as \( n \to \infty \).

For \( \alpha = 1 \), we have

\[ I_{1n} \leq C \sup_{z} \left| \hat{F}(z) - F(z) \right| \frac{1}{a_{n}} \int \int |\phi(x)| \left| k \left( \frac{x-y}{a_{n}} \right) \right| dF_{n}(x)dF_{n}(y) \]

\[ = CG \int |\phi(x)| \hat{f}(x)dF_{n}(x). \]
\[ I_{3n} = \left| \int \phi(x) \psi[F(x)] \hat{E}_n(x) dF_n(x) - \int \phi(x) \psi[F(x)] \hat{E}(x) dF(x) \right| \]

\[ = \frac{1}{n} \sum_{i=1}^{n} Y_i - EY_1 \rightarrow 0, \text{ w.p.l, as } n \rightarrow \infty, \]

by the SLLN and (xiv), establishing (2.14) for \( \alpha = 3 \). \( \square \)

Some conditions assumed by the above theorems seem difficult to verify and undesirable. The following two lemmas provide a set of sufficient conditions under which Conditions (v), (vi), (ix), (xi), (xii), and (xiv) are automatically satisfied.

**LEMMA 2.3.** (Nadaraya).

(2.17) \hspace{1cm} \text{Let } k \text{ be a function of bounded variation.}

(2.18) \hspace{1cm} \text{If, for any } c > 0, \sum_{n=1}^{\infty} \exp(-cna_n^2) < \infty,

then

(2.19) \hspace{1cm} \sup_{x} |\hat{f}(x) - \hat{E}(x)| \rightarrow 0, \text{ w.p.l, as } n \rightarrow \infty.

Furthermore, if

(2.20) \hspace{1cm} f \text{ is a uniformly continuous pdf,}
where \( c = 2\epsilon^2 \) and \( 0 < c_1 < \infty \). The last inequality in (2.22) is obtained by applying a result of Dvoretzky, Kiefer and Wolfowitz (1956). Thus, in view of the Borel-Cantelli lemma and (2.18), we have

\[
(2.23) \quad \sup_x |\hat{F}(x) - EF(x)| \to 0, \text{w.p.1, as } n \to \infty.
\]

Furthermore,

\[
(2.24) \quad \sup_x |EF(x) - F(x)| \leq \sup_x \int_{x^-}^x \sup_{u \leq x} \int_{(x-u)/a_n}^\infty k(z)dz f(u)du
\]

\[
+ \sup_x \int_x^\infty \sup_{u > x} \int_{-\infty}^{(x-u)/a_n} k(z)dz f(u)du.
\]

Let \( \gamma \) be an arbitrary small positive quantity. Then there exists \( M > 0 \) such that, for all \( n \geq M \), we have

\[
(2.25) \quad 1 - K\left(\frac{x-u}{a_n}\right) < \frac{1}{2}\gamma \quad \text{whenever } u < x, \text{ and}
\]

\[
(2.26) \quad K\left(\frac{x-u}{a_n}\right) < \frac{1}{2}\gamma \quad \text{whenever } u > x.
\]

Note that \( \gamma \) does not depend on \( x \) since \( K \) is a uniformly continuous df and thus for all \( x \) such that \( u < x, (u > x) \) the argument of \( K \) is
PROOF. It suffices to verify Conditions (iv), (ix) – (xiv). Now, (iv) is implied by (viii) and (2.20); (ix) is implied by (xii) which, in turn, is implied by (2.17) and (2.18); (x) is implied by (2.20); (xi) is implied by (2.17), (2.18), and (2.20); (xiii) is now replaced by (xv) in that \( \mu \) in (2.15) is replaced by \( \nu \sup f \); and, finally, (xiv) is implied by (2.20) and (viii). \( \square \)

3. Asymptotic Normality of the Estimate. In this section, asymptotic normality (AN) of \( n^{1/2}(\hat{\theta} - \theta) \) is established under some regularity conditions. This result enables us to obtain an approximate \((1 - \alpha)100\%\) confidence interval for the density functional \( \theta \). The method of proof is to find an appropriate sequence of i.i.d. random variables \( W_1, \ldots, W_n \), say, with \( E W_1^2 < \infty \); and then show that the asymptotic distribution of \( \hat{\theta} \) is the same as that of \( (1/n)\sum_{i=1}^{n} W_i \).

**Theorem 3.1.** Assume (iv) and the conditions of Theorem 2.2. Then, as \( n \to \infty \)

\[
(3.1) \quad n^{1/2}(\hat{\theta} - \theta) \sim \text{AN}(0, \sigma^2),
\]

where

\[
(3.2) \quad \sigma^2 = E[\phi(X)\psi[F(X)]f(X)]^2 - \theta^2.
\]
\[ \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} \{ \phi(X_i) \psi(\hat{F}(X_i)) \hat{f}(X_i) \}^2 - \hat{\theta}^2 \]

is a consistent estimate for \( \sigma^2 \) provided that \( \sigma^2 < \infty \).

4. Some Remarks. The functional \( \theta \) has many special cases that stem from studying the ARE's of various rank tests. Two examples of the two-sample problem are given here to illustrate the usefulness of the estimate \( \hat{\theta} \) obtained in (1.3). Let \( F \) and \( G \) be two univariate df's. First, consider testing \( H_0: F = G \) vs \( H_1: G(x) = F(x - \nu) \), where \( \nu \neq 0 \) is a location shift parameter. The ARE of the test using the celebrated Chernoff-Savage statistic \( T_N \) (see, e.g., Puri and Sen (1971), pp. 93-100) with respect to the t-test, when \( F \) is arbitrary and has finite variance \( \sigma^2 \), is given by (Puri and Sen (1971), p. 117)

\[ e_{T_N,t} = \sigma^2 \left[ \int J'(F(x)) f^2(x) dx \right]^2 / A^2. \]

where

\[ A^2 = \int_0^1 J^2(x) dx - \left[ \int_0^1 J(x) dx \right]^2. \]

Note that the dominant term in (4.1) is a special case of \( \theta \) with \( \phi(x) = 1 \) and \( \psi = J' \). A special case of the Chernoff-Savage statistic is the Mann-Whitney-Wilcoxon statistic for which the ARE with respect to the t-test becomes
REFERENCES


