King Fahd University of Petroleum & Minerals

DEPARTMENT OF MATHEMATICAL SCIENCES

Technical Report Series

TR 089

November 1986

Estimating Functionals of Probability Density Under Uniform Mixing Process

Khaled I. Abdul-Al

DHAHRAN 31261 ● SAUDI ARABIA ● www.kfupm.edu.sa/math/ ● E-mail: mathdept@kfupm.edu.sa

ESTIMATING FUNCTIONALS OF PROBABILITY DENSITY UNDER UNIFORM MIXING PROCESS

by

Khaled Abdulal
Department of Mathematical Sciences
University of Petroleum & Minerals
Dhahran, Saudi Arabia

Abstract

This paper is concerned with non parametric estimation of the distribution function $F(x)$, of the p-th order quantile, of the failure rate and of the functional $\int \gamma(x) f^2(x)dx$. Asymptotic properties estimators have been studied.

AMS Subject Classification: 62G05

Keywords: Density estimation, quantile, failure function, uniform mixing.

1. Introduction

Let $\{X_n\}$ be stationary uniform mixing process, then X_n's have the same probability density function (p.d.f.) $f(x)$ and the distribution function (d.f.) $F(x)$. Let $K(u)$ be a known symmetric p.d.f. satisfying the following condition

$$\sup_u K(u) < \infty \quad \text{and} \quad \lim_{|u| \to \infty} |u| K(u) = 0 \tag{1.1}$$

Also, let $\{a_n\}$ be a sequence of real numbers such that

$$a_n \to 0 \quad \text{as} \quad n \to \infty \tag{1.2}$$

Based on the first n observations an estimate of $f(x)$ is given by

$$f_n(x) = \frac{1}{n\, a_n} \sum_{j=1}^{n} K\left(\frac{x - X_j}{a_n}\right) \tag{1.3}$$

Among many applications of the density estimates, we shall single out four areas. First, we shall estimate the d.f. $F(x)$. As an approximation to $F(x)$ based on the first n observations, we take the statistic

$$F_n(x) = \int_{-\infty}^{\infty} f_n(t)dt \tag{1.4}$$

where $f_n(x)$ is defined by (1.3). In Section 2, consistency, uniform consistency, and asymptotic normality of $F_n(x)$ will be studied for uniform mixing sequences.

Second, we estimate the p-th order quantile of $F(x)$. For $0 < p < 1$, the p-th quantile of $F(x)$ is a root of the equation

$$F(\xi_p) = p \tag{1.5}$$

We assume that it is unique. An obvious estimate for it is the sample p-th quantile, that is, a root of the equation

$$F_n(\hat{\xi}_p) = p \tag{1.6}$$

In Section 3, weak consistency and asymptotic normality of $\hat{\xi}_p$ are established under the uniform mixing condition. The results of $F_n(x)$ and $\hat{\xi}_p$ are extensions of Nadarya [11]. Third, we shall study an estimate of the failure rate function $r(x)$. The failure rate is given by

$$r(x) = \frac{f(x)}{1 - F(x)} \tag{1.7}$$

for all x such that $F(x) < 1$. Using $f_n(x)$ and $F_n(x)$ given by (1.3) and (1.4), respectively, we estimate $r(x)$ by

$$r_n(x) = \frac{f_n(x)}{1 - F_n(x)} \tag{1.8}$$

In Section 4, we shall study the uniform consistency and the asymptotic normality of $r_n(x)$.

Finally, we shall study an estimate of the function

$$\theta(F) = \int \gamma(x)\ f^2(x)\ dx \tag{1.9}$$

where $\gamma(x)$ is measurable function. In Section 5, weak and strong consistency of the estimate

$$\theta_n(F) = \int \gamma(x)\ f_n^2(x)\ dx \tag{1.10}$$

will be established under the uniform mixing condition.

2. Preliminaries

Let $\{X_n\}$ be a stationary sequence of random variable defined on a probability space (Ω, B, P). For $a \leq b$ define $\sigma(a, b)$ as the σ-field generated by the random variables $X_a, \ldots, X_b$ and define $\sigma(a, \infty)$ as the σ-field generated by $X_a, X_{a+1}, \ldots,$.

We shall say that the sequence $\{X_n\}$ is uniform mixing if, for each $m (m \geq 1)$ and for each $n (n \geq 1)$, $A \in \sigma(1, m)$ and $B \in \sigma(m+n, \infty)$ together imply that

$$|P(AB) - P(A)P(B)| \leq \alpha(n)P(A) \tag{2.1}$$

where $\alpha(n)$, $n = 1,2,\ldots$ is a nonnegative function of integers such that

$$\lim_{n \to \infty} \alpha(n) = 0 \tag{2.2}$$

The following Lemmas will play a central role in this paper.

<u>Lemma 2.1</u>. Let $\{X_n\}$ be a uniform mixing stationary sequence and let the random variables η_1 and η_2 be measurable with respect to $\sigma(1, m)$ and $\sigma(m+n, \infty)$, respectively.

(a) $p > 1$, $q > 1$, are two real numbers such that $\frac{1}{p} + \frac{1}{q} = 1$, and if $E|\eta_1|^p < \infty$ and $E|\eta_2|^q < \infty$, then

$$|E[\eta_1\eta_2] - E[\eta_1]E[\eta_2]| \leq 2\{\alpha(n)E|\eta_1|^p\}^{1/p}\{E|\eta_2|^q\}^{1/q} \tag{2.3}$$

(b) If $|\eta_i| \leq c_i < \infty$ a.s., $i = 1,2$, then

$$|E[\eta_1\eta_2] - E[\eta_1]E[\eta_2]| \leq 2\alpha(n)c_1c_2 \tag{2.4}$$

<u>Proof</u>. Can be found in [5], pp.170-171.

<u>Lemma 2.2</u> (Asymptotic Unbiasedness). Let $\{X_n\}$ be a stationary sequence, and let $f_n(x)$ be given by (1.3). Suppose $K(y)$ satisfies condition (1.1) and the constants a_n satisfy (1.2). If $f(x)$ is continuous, then at all points x

$$\lim_{x \to \infty} E\, f_n(x) = f(x)$$

<u>Lemma 2.3</u>. Let $\{X_n\}$ be a stationary sequence of r.v.'s which satisfies the uniform mixing condition and let $f_n(x)$ be defined as in (1.3). Suppose that the following conditions hold:

(i) $f(x)$ is continuous and bounded,

(ii) $\sum_{j=1}^{\infty} \alpha^{\frac{1}{2}}(j) < \infty$,

(iii) $K(u)$ satisfies condition (1.1),

(iv) the constants a_n satisfy (1.2), and

(v) $f_j(x,y)$ [the joint density of X_1 and X_j, $j = 2,3,\ldots$ are continuous and bounded.] and $\sum_{j\neq 1} |f_j(x,y) - f(x)f(y)| < \infty \quad \forall x$ and y.

Then at all points x and y

$$\lim_{n\to\infty} na_n \operatorname{Cov}[f_n(x), f_n(y)] = \begin{cases} f(x)\int K^2(z)dz & \text{if } x = y \\ 0 & \text{if } x \neq y \end{cases}$$

Lemma 2.4. Assume $\{X_n\}$ is a sequence of stationary r.v.'s satisfying the uniform mixing condition and let $f_n(x)$ be defined as in (1.3). Assume that the following conditions hold:

(i) $f(x)$ is uniformly continuous,

(ii) the constants a_n satisfy (1.2), and

$$\lim_{n\to\infty} na_n^2 = \infty$$

(iii) $\int |k(t)|dt < \infty$, and

(iv) $\sum_{j=1}^{\infty} \alpha(j) < \infty$

Then

$$\sup_{-\infty<x<\infty} |f_n(x) - f(x)| \xrightarrow{P} 0 \qquad (\text{as } n \to \infty)$$

Lemma 2.5. Assume that $\{X_n\}$ is a sequence of stationary r.v.'s which satisfies the uniform mixing condition and let $f_n(x)$ be defined as in (1.3). Suppose that the following conditions hold:

(i) $f(x)$ is uniformly continuous,

(ii) The constants a_n satisfy (1.2), and

$$\sum_{n=1}^{\infty} \frac{1}{(na_n^2)^2} < \infty ,$$

(iii) $\int |k(t)| dt < \infty$, and

(iv) $\sum_{j=1}^{\infty} \alpha^{\frac{1}{2}}(j) < \infty$

Then

$$\sup_{-\infty<x<\infty} |f_n(x) - f(x)| \xrightarrow{W.P.1} 0 \qquad (\text{as } n \to \infty)$$

Theorem 2.6. Assume that the conditions of Lemma 2.3 are satisfied. Suppose that the following conditions hold:

(i) $\lim_{n \to \infty} na_n = \infty$,

(ii) for any pair of sequence $m = m(n)$, $r = r(n)$, such that m, $r \to \infty$ as $n \to \infty$ but $m = o(n^{1/3} a_n^{-2/3})$, $r = o(m(n))$, and $\lim_{n \to \infty} n\, m^{-1}\, \alpha(r) = 0$, and

(iii) the joint density functions up to the fourth order are bounded and continuous.

If $f_n(x)$ is defined as in (1.3), then for any $\ell \geq 1$, and any points x_i, $i = 1,\dots,\ell$,

$$\sqrt{na_n}\, [f_n(x_i) - E\, f_n(x_i)] \qquad (i = 1,\dots,\ell)$$

(x_i are distinct) are asymptotically normal and independent with mean zero and variances given by

$$f(x_i) \int K^2(z)dz \qquad (i = 1,\ldots,\ell)$$

Proofs of the above results can be found in [1].

3. Estimation of the Distribution Function $F(x)$

The following theorem will play the central role in studying the asymptotic unbiasedness and the weak consistency of $F_n(x)$. This kind of consistency presented in the next theorem is known as the integrated consistency of $f_n(x)$.

Theorem 3.1. Assume the following conditions hold:

(i) $f(x)$ is continuous and bounded,

(ii) $\sum_{j=1}^{\infty} \alpha^{\frac{1}{2}}(j) < \infty$,

(iii) K satisfies condition (1.1),

(iv) the constants a_n satisfy (1.2) and

$$\lim_{n \to \infty} na_n = \infty, \quad \text{and}$$

(v) $f_j(x,y)$ [the joint density of X_1 and X_j, $j = 2,3,\ldots$] are continuous and bounded, then

$$\lim_{n \to \infty} E\int |f_n(x) - f(x)|dx = 0 \tag{3.1}$$

Proof. Note that from Lemma 2.2

$$E\ f_n(x) \to f(x) \qquad \text{as} \quad n \to \infty \quad \text{at all points} \quad x \tag{3.2}$$

and

$$\begin{aligned}\int E\ f_n(x)dx &= \iint \frac{1}{a_n} K\left(\frac{x-y}{a_n}\right) f(y)\,dy\,dx \\ &= \iint K(z)\ f(x - a_n z)\,dz\,dx = 1\end{aligned} \tag{3.3}$$

Also $\qquad |E\ f_n(x) - f(x)| \leq E\ f_n(x) + f(x)$

which is integrable and converges to the integrable function $2f(x)$ as $n \to \infty$ by (2.2) and (2.3). Hence by Lebesgue's Dominated Convergence Theorem (LDCT), we have

$$\int |E\ f_n(x) - f(x)|dx \to 0 \qquad \text{as} \quad n \to \infty \tag{3.4}$$

Next, note that by Schwartz's inequality and Lemma 2.3, we have

$$E|f_n(x) - E\ f_n(x)| \leq [\text{Var}\ f_n(x)]^{\frac{1}{2}} \to 0 \quad \text{as} \quad n \to \infty$$

Also note that

$$E|f_n(x) - E\ f_n(x)| \leq 2\,E\,f_n(x)$$

which is integrable and converges to integrable function $2f(x)$ as $n \to \infty$ by (3.2) and (3.3).

From Fubini's Theorem and LDCT, we have

$$E\int |f_n(x) - E\, f_n(x)|dx = \int E|f_n(x) - E\, f_n(x)|dx \to 0 \quad \text{as} \quad n \to \infty \qquad (3.5)$$

Now, from (3.4) and (3.5), and the inequality

$$E\int |f_n(x) - f(x)|dx \le E\int |f_n(x) - E\, f_n(x)|dx + \int |E\, f_n(x) - f(x)|dx$$

the theorem is true.

Next, we discuss the asymptotic unbiasedness and the weak consistency of $F_n(x)$.

<u>Theorem 3.2</u>. Let $\{X_n\}$ be a sequence of stationary r.v.'s satisfying the uniform mixing condition and let $F_n(x)$ be defined as in (1.4). Assume that the following conditions hold:

(i) $f(x)$ is continuous and bounded,

(ii) $\sum_{j=1}^{\infty} \alpha(j) < \infty$,

(iii) $K(y)$ satisfies (1.1), and

(iv) the constants a_n satisfy (1.2).

Then at all points x

$$\lim_{n \to \infty} E\, F_n(x) = F(x), \tag{3.6}$$

$$\lim_{n} \operatorname{Var} F_n(x) = 0 \tag{3.7}$$

and

$$F_n(x) \xrightarrow{P} F(x) \qquad (\text{as } n \to \infty) \tag{3.8}$$

Proof. The first part of the theorem follows immediately from the following inequality and (2.4).

$$\begin{aligned} |E\, F_n(x) - F(x)| &\le \int_{-\infty}^{x} |E\, f_n(t) - f(t)|dt \\ &\le \int |E\, f_n(t) - f(t)|dt \to 0 \quad \text{as} \quad n \to \infty \end{aligned}$$

Next, we compute the variance of $F_n(x)$. By stationarity

$$\begin{aligned} \operatorname{Var} F_n(x) &= \frac{1}{na_n^2} \operatorname{Var}\Big[\int_{-\infty}^{x} K\Big(\frac{t - X_1}{a_n}\Big)dt\Big] \\ &\quad + \frac{1}{n^2a_n^2} \sum_{j=2}^{n} (n-j+1) \operatorname{Cov}\Big[\int_{-\infty}^{x} K\Big(\frac{t - X_1}{a_n}\Big)dt,\ \int_{-\infty}^{x} K\Big(\frac{t - X_j}{a_n}\Big)dt\Big] \\ &= I_{n_1} + I_{n_2} \end{aligned} \tag{3.9}$$

Consider I_{n_1} then

$$\frac{1}{a_n^2}\,[\mathrm{Var}[\int_{-\infty}^{x} K(\frac{t - X_1}{a_n})dt] = \int [\int_{-\infty}^{\frac{x-y}{a_n}} K(t)dt]^2 f(y)dy$$

$$- [(\int_{-\infty}^{\frac{x-y}{a_n}} K(t)dt)\ f(y)dy]^2$$

Since $[\int_{-\infty}^{\frac{x-y}{a_n}} K(t)dt]^r f(y)$ converges to $I(x - y)f(y)$ which is an integrable function, where $r = 1,2,$ and

$$I(x - y) = \begin{cases} 0 & \text{if} \quad x < y \\ 1 & \text{if} \quad x \geq y \end{cases}$$

then by Lebesgue Dominated Convergence Theorem, we have

$$I_{n_1} \cong \frac{1}{n} \int I(x - y)f(y)dy - \frac{1}{n} [\int I(x - y)f(y)dy]^2$$

$$= \frac{F(x)(1 - F(x))}{n} \tag{3.10}$$

Consider I_{n_1} then

$$\frac{1}{a_n^2} \mathrm{Cov}[\int_{-\infty}^{x} K(\frac{t - X_1}{a_n})dt,\ \int_{-\infty}^{x} K(\frac{t - X_j}{a_n})dt]$$

$$= \mathrm{Cov}[K^*(\frac{x - X_1}{a_n}),\ K^*(\frac{x - X_j}{a_n})] \tag{3.11}$$

where $K^*(x) = \int_{-\infty}^{x} K(t)dt$

But $\left| K^*(\frac{x - X_j}{a_n}) \right| \leq 1$

Then from Lemma 2.1, I_{n_2} without $1/n$ will be bounded, i.e.,

$$\frac{1}{n}\sum_{j=2}^{n}(n-j+1)\,\mathrm{Cov}[K^*(\frac{x-X_1}{a_n}),\ K^*(\frac{x-X_j}{a_n})]$$

$$\leq \sum_{j=2}^{n}\mathrm{Cov}\left|[K^*(\frac{x-X_1}{a_n}),\ K^*(\frac{x-X_j}{a_n})]\right| \tag{3.12}$$

$$\leq 2\sum_{j=1}^{\infty}\alpha(j)$$

From (3.9) - (3.12), we have

$$\lim_{n\to\infty}\mathrm{Var}[F_n(x)] = 0 \qquad \text{at all points } x$$

and

$$\lim_{n\to\infty} n\,\mathrm{Var}\,F_n(x) < \infty \qquad \text{at all points } x$$

From (3.6) and (3.7), it follows that

$$E[F_n(x) - F(x)]^2 = \mathrm{Var}\,F_n(x) + [EF_n(x) - F(x)]^2 \to 0 \quad \text{as} \quad n\to\infty \tag{3.13}$$

From (3.13) and Markov's inequality, it follows that

$$F_n(x) \xrightarrow{P} F(x) \qquad \text{at all points } x \quad (n\to\infty)$$

Next, we discuss the uniform convergence of $F_n(x)$.

<u>Theorem 3.3</u>. Assume that the conditions of Lemma 2.5 hold. Then

$$\sup_{-\infty<x<\infty} |F_n(x) - F(x)| \xrightarrow{W.P.1} F(x) \qquad \text{as} \quad n \to \infty \tag{3.14}$$

<u>Proof</u>. It is divided into three steps:

(1) We have to show that

$$f_n(x) \xrightarrow{W.P.1} f(x) \qquad \text{at all points} \quad x \quad \text{as} \quad n \to \infty \tag{3.15}$$

This follows from Lemma 2.5 and the inequality

$$|f_n(x) - f(x)| \leq \sup_{-\infty<x<\infty} |f_n(x) - f(x)|$$

(2) We have to show that

$$F_n(x) \xrightarrow{W.P.1} F(x) \qquad \text{at all points} \quad x \quad \text{as} \quad n \to \infty \tag{3.16}$$

It follows from (3.15) and Lebesgue Dominated Convergence Theorem

$$|F_n(x) - F(x)| = \left|\int_{-\infty}^{x} [f_n(t) - f(t)]dt\right| \leq \int_{-\infty}^{x} |f_n(t) - f(t)|dt \tag{3.17}$$

$$\leq \int |f_n(t) - f(t)|dt \to 0 \qquad \text{as} \quad n \to \infty$$

(3) Use (3.17) and proceed exactly as in the proof of Glivenko-Cantelli Lemma (See Loéve, [10], p.20).

The remainder of this section will be devoted to the study of the asymptotic normality of the estimate $F_n(x)$.

Theorem 3.4. Assume conditions of Theorem 3.2 are satisfied. Then if $\operatorname{Var} F_n(x) > 0$

$$\frac{[F_n(x) - E\,F_n(x)]}{\sqrt{\operatorname{Var} F_n(x)}} \xrightarrow{D} N(0,\ 1) \qquad \text{as} \quad n \to \infty \tag{3.18}$$

Proof. Let

$$z_j = K^*\left(\frac{x - X_j}{a_n}\right) - E\,K^*\left(\frac{x - X_j}{a_n}\right) \tag{3.19}$$

where $K^*(y)$ is defined as in (3.11)

Notice that

$$P[|z_j| \leq 1] = 1 \qquad \text{for} \quad j = 1,2,\ldots,n \tag{3.20}$$

From (3.19) and Theorem 18.5.4 of Ibragimov and Linnik [9], it follows that (3.18) is true.

4. Estimation of the Distribution Quantiles

Let ξ_p and $\hat{\xi}_p$ be defined as in (1.5) and (1.6), respectively, where $0 < p < 1$. We assume ξ_p to be unique. In the present section two properties of $\hat{\xi}_p$ will be established. Namely, consistency in the probability sense, and asymptotic normality. The first of these results is the following theorem.

Theorem 4.1. Assume that the conditions of Lemma 2.5 hold.

Then

$$\hat{\xi}_p \xrightarrow{P} \xi_p \quad \text{as} \quad n \to \infty \qquad 0 < p < 1 \tag{4.1}$$

<u>Proof</u>. For $\varepsilon > 0$, define $\delta(\varepsilon)$ by

$$\delta(\varepsilon) = \min\{F(\xi_p + \varepsilon) - F(\xi_p),\ F(\xi_p) - F(\xi_p - \varepsilon)\} \tag{4.2}$$

Then $\delta(\varepsilon)$ is positive because of the uniqueness of ξ_p. Next

$$|F(\hat{\xi}_p) - F(\xi_p)| = |F(\hat{\xi}_p) - F_n(\hat{\xi}_p)| \tag{4.3}$$

$$\leq \sup_{-\infty<X<\infty} |F_n(x) - F(x)|$$

While from the definition of $\delta(\varepsilon)$ it follows that

$$\{|\hat{\xi}_p - \xi_p| > \varepsilon\} \subseteq \{|F(\hat{\xi}_p) - F(\xi_p)| > \delta(\varepsilon)\} \tag{4.4}$$

Therefore,

$$P\{|\hat{\xi}_p - \xi_p| > \varepsilon\} \leq P\{|F(\hat{\xi}_p) - F(\xi_p)| > \delta(\varepsilon)\} \tag{4.5}$$

$$\leq P\{\sup_{-\infty<X<\infty} |F_n(x) - F(x)| > \delta(\varepsilon)\}$$

where $\sup_{-\infty<X<\infty} |F_n(x) - F(x)| \to 0$ in probability by Theorem 3.3. From (4.5), it then follows that

$$P\{|\hat{\xi}_p - \xi_p| > \varepsilon\} \to 0 \quad \text{as} \quad n \to \infty \tag{4.6}$$

Thus, the consistency of the estimate $\hat{\xi}_p$ has been proved.

The proof of the asymptotic normality is much more involved. Some preliminary results will facilitate it. Applying Taylor's expansion to $F_n(\hat{\xi}_p)$, we get

$$P = F(\xi_p) = F_n(\hat{\xi}_p) = F_n(\xi_p) + (\hat{\xi}_p - \xi_p)f_n(\xi) \tag{4.7}$$

where ξ is some random point between $\hat{\xi}_p$ and ξ_p. From (4.7) we obtain

$$\begin{aligned}\frac{f(\xi_p)(\hat{\xi}_p - \xi_p)}{\sqrt{\operatorname{Var} F_n(\hat{\xi}_p)}} = &- \left\{\frac{F_n(\xi_p) - E\, F_n(\xi_p)}{\sqrt{\operatorname{Var} F_n(\hat{\xi}_p)}}\ \frac{f(\xi_p)}{f_n(\xi)}\right\} \\ &- \left\{\frac{E\, F_n(\xi_p) - F(\xi_p)}{\sqrt{\operatorname{Var} F_n^*(\hat{\xi}_p)}}\ \frac{f(\xi_p)}{f_n(\xi)}\right\}\end{aligned} \tag{4.8}$$

If we can show the following:

(i) $f_n(\xi) \to f(\xi_p)$ in probability as $n \to \infty$

(ii) $\left\{\frac{F_n(\xi_p) - E\, F_n(\xi_p)}{\sqrt{\operatorname{Var} F_n(\hat{\xi}_p)}}\ \frac{f(\xi_p)}{f_n(\xi)}\right\} \to 0$ in probability as $n \to \infty$

then

$$\frac{f(\xi_p)(\hat{\xi}_p - \xi_p)}{\sqrt{\operatorname{Var} F_n(\hat{\xi}_p)}} \quad \text{and} \quad \frac{F_n(\xi_p) - E\, F_n(\xi_p)}{\sqrt{\operatorname{Var} F_n(\hat{\xi}_p)}}$$

will have the same limiting distribution.

Theorem 4.2. Assume that the conditions of Lemma 2.4 and Theorem 4.1 are satisfied. Assume that the following conditions hold:

(i) $K(y)$ is symmetric and $\int y^2 K(y)dy < \infty$, (4.10)

(ii) $na_n^4 \to 0$ as $n \to \infty$, and (4.11)

(iii) $F(x)$ has a bounded second derivative.

Then

$$\frac{f(\xi_p)(\hat{\xi}_p - \xi_p)}{\text{Var } F_n(\hat{\xi}_p)} \xrightarrow{D} N(0, 1) \qquad \text{as } n \to \infty$$

<u>Proof</u>. First, we have to show that

$$f_n(\xi) \to f(\xi_p) \quad \text{in probability as } n \to \infty \tag{4.12}$$

$$|f_n(\xi) - f(\xi_p)| \leq |f_n(\xi) - f(\xi)| + |f(\xi) - f(\xi_p)|$$

$$\leq \sup_{-\infty<x<\infty} |f_n(x) - f(x)| + |f(\xi) - f(\xi_p)| \tag{4.13}$$

Now, from Lemma 2.4, $\sup_{-\infty<x<\infty} |f_n(x) - f(x)| \to 0$ in probability and also $|f(\xi) - f(\xi_p)| \to 0$ in probability because $f(x)$ is continuous and $\xi \to \xi_p$ in probability, hence (4.12) follows.

Next, we show that

$$\sqrt{n}\,[E\,F_n(x) - F(x)] \to 0 \quad \text{as } n \to \infty \tag{4.14}$$

Using Taylor's expansion, we have

$$\sqrt{n}\,[E\,F_n(x) - F(x)] = \sqrt{n}\,[\int\int_{-\infty}^{x} \frac{1}{a_n} K(\frac{t-y}{a_n})\, f(y)\, dy\, dt - F(x)]$$

$$= \sqrt{n}\ \{\int [F(x - a_n y) - F(x)]K(y)dy\} \tag{4.15}$$

$$= \sqrt{n}\ \{\int [a_n y\ F^{(1)}(x)\ K(y) + \frac{a_n^2}{2} F^{(2)}(x^*)y^2K(y)]dy\}$$

It follows from (4.15) and conditions (i) to (iii) that

$$\sqrt{n}\ [E\ F_n(x) - F(x)] = \sqrt{n}\ \frac{a_n^2}{2}\ \{F^{(2)}(x) \int y^2K(y)dy\} \to 0 \quad \text{as} \quad n \to \infty$$

We conclude from (4.12), (4.14) and (3.7) that

$$\frac{[E\ F_n(\xi_p) - F(\xi_p)]f(\xi_p)}{\sqrt{\text{Var } F_n(\hat{\xi}_p)}\ f_n(\xi)} \longrightarrow 0 \quad \text{in probability} \tag{4.16}$$

From (4.8), (4.12), (4.16), and Theorem (3.4), it then follows that

$$\frac{f(\xi_p)(\hat{\xi}_p - \xi)}{\sqrt{\text{Var } F_n(\hat{\xi}_p)}} \xrightarrow{D} N(0,\ 1)$$

5. Estimation of Failure Rate Function

Results of Sections 2 and 3 will be utilized here to study the asymptotic properties of $r_n(x)$ proposed in (1.8). In particular, the pointwise consistency, strong uniform consistency and asymptotic normality of $r_n(x)$ will be established.

Theorem 5.1. Assume that the conditions of Lemma 2.5 hold. Then at all points x

$$r_n(x) \xrightarrow{W.P.1} r(x) \qquad \text{as} \quad n \to \infty \tag{5.1}$$

Proof. This follows immediately from (3.15) and (3.16) and using a well-known convergence theorem see Cramer [6], p.254.

Theorem 5.2. Assume that $\{X_n\}$ is a sequence of stationary r.v.'s which satisfies the uniform mixing condition and let $r_n(x)$ be defined as in (1.8). Suppose that the following conditions hold:

(i) $f(x)$ is uniformly continuous on $[a, b]$, $-\infty \le a < b < \infty$, and $\bar{F}(b) = [1 - F(b)] > 0$, and

(ii) conditions (ii) - (iv) of Lemma 2.5.

Then

$$\sup_{a \le x \le b} |r_n(x) - r(x)| \xrightarrow{W.P.1} 0 \qquad \text{as} \quad n \to \infty \tag{5.2}$$

Proof. Define $\bar{F}_n(x) = 1 - F_n(x)$.
Note that $\bar{F}(b) \le \bar{F}(x) \le \bar{F}(a)$ and $\bar{F}_n(b) \le \bar{F}_n(x) \le \bar{F}_n(a)$ for all $x \in [a,b]$. Now,

$$\sup_{a \le x \le b} |r_n(x) - r(x)| = \sup_{a \le x \le b} \left| \frac{f_n(x)}{\bar{F}_n(x)} - \frac{f(x)}{\bar{F}(x)} \right|$$

$$= \sup_{a \le x \le b} \left| \frac{f_n(x)}{\bar{F}_n(x)} - \frac{f_n(x)}{\bar{F}(x)} + \frac{f_n(x)}{\bar{F}(x)} - \frac{f(x)}{\bar{F}(x)} \right| \tag{5.3}$$

$$\le \sup_{a \le x \le b} f_n(x) \left| \frac{1}{\bar{F}_n(x)} - \frac{1}{\bar{F}(x)} \right| + \sup_{a \le x \le b} |f_n(x) - f(x)| \left| \frac{1}{\bar{F}(x)} \right|$$

$$= I_{n_1} + I_{n_2}$$

where

$$I_{n_1} = \sup_{a\le x\le b} f_n(x)|F_n(x) - F(x)|\left|\frac{1}{\bar{F}_n(x)\bar{F}(x)}\right|$$

$$\le \frac{1}{\bar{F}_n(b)\bar{F}(b)} \sup_{a\le x\le b} f_n(x) \sup_{a\le x\le b} |F_n(x) - F(x)| \tag{5.4}$$

which converges to 0 W.P.1 as $n \to \infty$ by Theorem 3.3 and the fact that $\sup_{-\infty<x<\infty} |K(x)| < \infty$. Finally, for I_{n_2} note that

$$I_{n_2} \le \frac{1}{\bar{F}(b)} \sup_{a\le x\le b} |f_n(x) - f(x)| \tag{5.5}$$

which converges to 0 W.P.1 as $n \to \infty$ by Lemma 2.5. The proof of the theorem is now complete.

In the next theorem we will establish the asymptotic normality of $r_n(x)$.

<u>Theorem 5.3</u>. Assume that the conditions of Theorem 2.6 are satisfied. If the following conditions hold:

(i) $K(y)$ is symmetric and $\int y^2K(y)dy < \infty$

(ii) $na_n^5 \to \infty$ as $n \to \infty$

(iii) $f(x)$ has a bounded second derivative, and

(iv) $\int y^2K(y)dy < \infty$, then

$$Z_n = \sqrt{na_n}\,[r_n(x) - r(x)\,\bar{F}(x)] \tag{5.6}$$

is asymptotically normal with mean zero and variance given by

$$f(x)\int K^2(z)dz \tag{5.7}$$

<u>Proof</u>. Following the proof of Theorem 7 of Watson and Leadbetter [13], we get

$$Z_n = \{\bar{F}(x)[r_n(x) - \frac{E\ f_n(x)}{\bar{F}_n(x)}] - \frac{E\ f_n(x)}{\bar{F}_n(x)}[\bar{F}_n(x) - \bar{F}(x)]$$

$$+ [E\ f_n(x) - f(x)]\}\sqrt{na_n} \tag{5.8}$$

$$= I_{n_1} + I_{n_2} + I_{n_3}$$

By Theorem 2.6, we have

$$\sqrt{na_n}\ [f_n(x) - E\ f_n(x)] \tag{5.9}$$

converges to normal distribution, and the fact that

$$F_n(x) \xrightarrow{P} F(x) \quad \text{as} \quad n \to \infty \tag{5.10}$$

Combining (5.9) and (5.10) and using a well-known convergence theorem [see Cramer [6]), p.254], we get that

$$I_{n_1} \xrightarrow{D} N(0,\ f(x)\int K^2(z)dz) \tag{5.11}$$

From Theorem 3.2, we have

$$[F_n(x) - F(x)] = O_p(n^{-\frac{1}{2}}) \tag{5.12}$$

where $O_p(\cdot)$ denotes "bounded in probability." From (5.12) and Lemma 2.2, we get

$$I_{n_2} \xrightarrow{P} 0 \quad \text{as} \quad n \to \infty \tag{5.13}$$

Consider I_{n_3}, then from conditions (i) - (iii)

$$I_{n_3} = \sqrt{na_n}\,[E\,f_n(x) - f(x)]$$

$$= \sqrt{na_n}\,[\int \frac{1}{a_n} K(\frac{x-y}{a_n})\; f(y)dy - f(x)]$$

$$= \sqrt{na_n}\,[\int [f(x - a_n y) - f(x)]K(y)dy$$

$$= \sqrt{na_n^5}\,[f^{(2)}(x^*) \int y^2 K(y)dy] \to 0 \quad \text{as} \quad n \to \infty$$

where x^* is between $x - a_n y$ and x.

6. Estimation of the Functional $\int \gamma(x) f^2(x)dx$

The functional $\theta(F) = \int \gamma(x) f^2(x)dx$ plays an important role in the studies of asymptotic relative efficiencies for various rank test statistics. For practical purposes, it is desirable to provide a consistent estimate for $\theta(F)$.

In this section we shall show that the estimate

$$\theta_n(F) = \int \gamma(x) f_n^2(x)dx \tag{6.1}$$

is asymptotically unbiased, weakly and strongly consistent.

Theorem 6.1. Let $\{X_j\}$ be a stationary sequence of r.v.'s satisfying the uniform mixing condition. Assume that the following condition hold:

(i) $K(x)$ satisfies (1.1)

(ii) $f(x)$ is continuous,

(iii) $\sum_{j=1}^{\alpha} \alpha^{\frac{1}{2}}(x) < \infty$, and

(iv) the constants a_n satisfy (1.2) and $\lim_{n \to \infty} na_n = \infty$

(v) $\int \gamma(x)f(x)dx < \infty$, $\int \gamma(x)K^2(x)dx < \infty$

Then

$$\operatorname*{Lim}_{n \to \infty} E\,\theta_n(F) = \theta(F) \tag{6.2}$$

Proof. By stationarity, we have

$$E\theta_n(F) = E\int \gamma(x)f_n^2(x)dx = \frac{1}{na_n^2} E\Big[\int \gamma(x)K^2\Big(\frac{x - X_1}{a_n}\Big)dx\Big] \tag{6.3}$$

$$+ \frac{1}{n^2a_n^2}\sum_{j=2}^{n}(n-j+1)\; E\Big[\int \gamma(x)K\Big(\frac{x-X_1}{a_n}\Big)K\Big(\frac{x-X_j}{a_n}\Big)dx\Big]$$

$$= I_{n_1} + I_{n_2}$$

By Fubini's Theorem, we have

$$I_{n_1} = \frac{1}{na_n^2}\iint \gamma(x)K^2(\frac{x-y}{a_n})f(y)dydx \tag{6.4}$$

$$= \frac{1}{na_n}\iint \gamma(x)K^2(z)f(x - a_n z)dxdz$$

Applying Lemma 3.2 in [1], we will have

$$\int K^2(z)f(x - a_n z)dz \to f(x)\int K^2(z)dz \tag{6.5}$$

It follows from (6.4) and (6.5) and the fact that $\int\gamma(x)f(x)dx$ and $\int K^2(z)dz$ are finite,

$$In_1 \cong \frac{1}{na_n}\iint \gamma(x)f(x)K^2(z)dzdx \tag{6.6}$$

But $na_n \to \infty$ as $n \to \infty$, so

$$I_{n_1} \to 0 \quad \text{as} \quad n \to \infty. \tag{6.7}$$

Consider I_{n_2}, then

$$I_{n_2} = \frac{2}{n^2a_n^2}\sum_{j=2}^{n}(n - j + 1)[\int \gamma(x)EK(\frac{x - X_1}{a_n})K(\frac{x - X_j}{a_n}) \tag{6.8}$$

$$- \int \gamma(x)E^2K(\frac{x - X_1}{a_n})dx] + \frac{n-1}{na_n^2}\int \gamma(x)E\ K(\frac{x - X_1}{a_n})dx$$

$$= J_{n_1} + J_{n_2}$$

$$J_{n_1} \le \frac{2}{na_n^2} \sum_{j=2}^{n} \int \gamma(x) \left| EK(\frac{x - X}{a_n})K(\frac{x - X_j}{a_n}) - EK(\frac{x - X}{a_n})EK(\frac{x - X_j}{a_n}) \right| dx$$

$$\le \frac{2}{na_n^2} \sum_{j=2}^{n} \int 2\gamma(x) EK^2(\frac{x - X_1}{a_n}) \alpha^{\frac{1}{2}}(j - 1) dx \tag{6.9}$$

$$\cong \frac{4}{na_n} \sum_{j=2}^{n} \alpha^{\frac{1}{2}}(j) [\iint \gamma(x) f(x) K^2(z) dx dz] \to 0 \quad \text{as} \quad n \to \infty$$

The second inequality by Lemma 2.1, the asymptotic equality by Lemma 3.2 in [1], and the limit converge to 0 because $na_n \to \infty$ as $n \to \infty$.

$$J_{n_2} = \frac{n - 1}{na_n^2} \int \gamma(x) E^2 K(\frac{x - X_1}{a_n}) dx \tag{6.10}$$

$$= \frac{n - 1}{n} \int \gamma(x) E \frac{1}{a_n} K(\frac{x - X_1}{a_n})]^2 dx$$

It follows from (6.10), Lemma 2.2, and LDCT, that

$$J_{n_2} \cong \frac{n - 1}{n} \int \gamma(x) f^2(x) dx \to \int \gamma(x) f^2(x) dx \quad \text{as} \quad n \to \infty \tag{6.11}$$

From (6.7), (6.8), (6.9), and (6.11), it follows that

$$\lim_{n \to \infty} E\theta_n(F) = \theta(F).$$

Next we prove the weak consistency.

Theorem 6.2. Assume the conditions of Theorem 6.1 are satisfied, then at all points x

$$\theta_n(F) \xrightarrow{P} \theta(F) \qquad \text{as} \quad n \to \infty \tag{6.12}$$

<u>Proof</u>. It is enough to show that

$$E|\theta_n(F) - \theta(F)| \to 0 \quad \text{as} \quad n \to \infty \tag{6.13}$$

Now,
$$E|\theta_n(F) - \theta(F)| \le E|\theta_n(F) - E\theta_n(F)| + |E\theta_n(F) - \theta(F)| \tag{6.14}$$

$$= I_{n_1} + I_{n_2}$$

From Theorem 6.1, we have

$$I_{n_2} \to 0 \quad \text{as} \quad n \to \infty \tag{6.15}$$

Consider I_{n_1}, then

$$I_{n_1} \le E\left|\int \gamma(x) f_n^2(x)dx - \int \gamma(x) E f_n^2(x)dx\right|$$

$$\le E\left|\int \gamma(x) f_n^2(x)dx - \int \gamma(x) E^2 f_n(x)dx\right| + \left|\int \gamma(x) E^2 f_n(x)\right.$$

$$\left. - \int \gamma(x) E f_n^2(x)dx\right| = J_{n_1} + J_{n_2}$$

$$J_{n_2} \to 0 \quad \text{as} \quad n \to \infty$$

because $\int \gamma(x) E^2 f_n(x)dx \to \int \gamma(x) f^2(x)dx$ by Lemma 3.2 and LDCT.

Now,

$$J_{n_1} \le \int |\gamma(x)| E|f_n^2(x) - E^2 f_n(x)|dx \tag{6.17}$$

$$\leq \int |\gamma(x)| E^{\frac{1}{2}} |f_n(x) + Ef_n(x)|^2 E^{\frac{1}{2}} |f_n(x) + Ef_n(x)|^2 dx$$

Since $\gamma(x)E^{\frac{1}{2}}|f_n(x) + Ef_n(x)|^2$ converges to $2f^2(x)\gamma(x)$ which is an integrable function, and $E^{\frac{1}{2}}|f_n(x) - Ef_n(x)|^2 = \sqrt{\mathrm{Var}\ f_n(x)}$ converges to 0 by Lemma 2.3, then

$$J_{n_1} \to 0 \quad \text{as} \quad n \to \infty \tag{6.18}$$

Combining (6.15), (6.16), we get (6.13). Using (6.13) and Markov's inequality, we get

$$\theta_n(F) \xrightarrow{P} \theta(F) \qquad \text{as} \quad n \to \infty$$

<u>Theorem 6.3</u>. Assume that the conditions of Lemma 2.5 are satisfied. If $\int|\gamma(x)|f(x)dx < \infty$ and $|\gamma(x)f(x)| < \infty$, then

$$\theta_n(F) \xrightarrow{W.P.1} \theta(F) \qquad \text{as} \quad n \to \infty \tag{6.19}$$

<u>Proof</u>.

$$|\theta_n(F) - \theta(F)| \leq |\theta_n(F) - \int\gamma(x)E^2f_n(x)dx| + |\int\gamma(x)E^2f_n(x) - \theta(F)|$$

$$= I_{n_1} + I_{n_2}$$

By Lemma 2.2 in [1], we have

$$I_{n_2} \to 0 \qquad \text{as} \quad n \to \infty \tag{6.20}$$

Consider I_{n_1}, then

$$I_{n_1} = \left|\int \gamma(x) f_n^2(x)dx - \int \gamma(x) E^2 f_n(x)dx\right| \qquad (6.21)$$

$$\leq \int |\gamma(x)||f_n(x) + Ef_n(x)||f_n(x) - Ef_n(x)|dx$$

$$\leq \sup_{-\infty<x<\infty} |f_n(x) - Ef_n(x)| \int |\gamma(x)||f_n(x) + Ef_n(x)|dx$$

But since $|\gamma(x)|(f_n(x) + Ef_n(x))$ converges with probability one $2|\gamma(x)f(x)|$, then by LDCT

$$\int |\gamma(x)|(f_n(x) + Ef_n(x))dx \xrightarrow{\text{W.P.1}} 2\int |\gamma(x)| f(x)dx \qquad (6.22)$$

It follows from (6.22) and Lemma 2.3, we have

$$I_{n_1} \xrightarrow{\text{W.P.1}} 0 \qquad \text{as } n \to \infty \qquad (6.23)$$

Combining (6.20) and (6.23), we get (6.19).

[1] Abdulal, K.I. and M.M. Siddiqui (1986). Density Estimation for Uniform Mixing Process. Bulletin of Informatics and Cybernetics, 22, pp.79-93.

[2] Ahmad, I.A. (1976). On the Asymptotic Properties of an Estimate of a Functional of a Probability Density. Scand. Actuarial J., 4, 176-181.

[3] Ahmad, I.A. and Lin, P.E. (1983). Consistency of a Nonparametric Estimation of a Density Functional. Metrika, 30, 21-29.

[4] Beans, S.J. and Tsokos, C.P. (1980). Developments in Nonparametric Density Estimation. Int. St. RVW, 48, 267-287.

[5] Billingsley, P. (1968). Convergence of Probability Measures. Wiley and Sons, New York.

[6] Cramer, H. (1946). Mathematical Methods of Statistics. Princeton University Press, Princeton, N.J.

[7] Doob, J. (1953). Stochastic Processes. Wiley and Sons, New York.

[8] Ibragimov, I.A. (1962). Some Limit Theorems for Stationary Processes. Theor. Probability Appl., 7, 349-382.

[9] Ibragimov, I.A. and Linnik, Yu J. (1971). Independent and Stationary Dependent Random Variables. Walters-Noordhoff Publishing, Groningen, Netherlands.

[10] Loeve, M. (1960). Probability Theory. Van Nostrand, Princeton.

[11] Nadaraya, E.A. (1964). Some New Estimates for Distribution Functions. Theor Probability Appl., 9, 497-500.

[12] Parzen, E. (1962). On the Estimation of a Probability Density Function and the Mode. Ann. Math Stat., 33, 1065-1076.

[13] Watson, G.S. and Leadbetter, M.R. (1964a). Hazard Analysis II. Sankhya Series A., 24, 101-116.

[14] Watson, G.S. and Leadbetter, M.R. (1964b). Hazard Analysis I. Biometrika, 51, 175-184.