THE UNIVERSITY OF CALGARY

NONPARAMETRIC ESTIMATION OF A HAZARD RATE

by

CHI CHUEN SO

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE

DEGREE OF

MASTER OF SCIENCE

DEPARTMENT OF

MATHEMATICS AND STATISTICS

CALGARY, ALBERTA

FEBRUARY, 1990

ISBN 0-315-61756-X

THE UNIVERSITY OF CALGARY
FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled, *"Nonparametric estimation of a hazard rate"*, submitted by Chi Chuen So in partial fulfillment of the requirements for the degree of Master of Science.

M. D. Burke

Supervisor, Dr. M.D. Burke,
Department of Mathematics and Statistics

David Paul Johnson

Dr. D.P. Johnson, Department of Mathematics and Statistics

J K Wani

Dr. J.K. Wani, Department of Mathematics and Statistics

Lawrence Harder

Dr. L.D. Harder, Department of Biological Sciences

DATE: 1990: 03: 15

ABSTRACT

In many reliability studies, the hazard rate h(x) is of prime importance. Therefore, the hazard rate estimation has gained considerable interest in statistical literature. Three significant estimators have been proposed and studied by Watson and Leadbetter (1963) and Rice and Rosenblatt (1976). The objective here is to provide a detailed and extensive analysis of these estimators.

In Chapter II, the bias and asymptotic unbiasedness of these three estimators are scrutinized thoroughly. Chapter III deals with the asymptotic equivalence, normality and global deviation. In Chapter IV, a numerical example is given to test and demonstrate the goodness of performance and the asymptotic properties of the estimators. Finally, I generalize the idea of these estimators to give an estimate of hazard rate for censored data by adapting the reduced sample technique in Chapter V. A simple experiment was also done to indicate how censorship strongly affects the estimate under this reduced sample technique.

ACKNOWLEDGEMENT

I wish to express my sincerest gratitude to my supervisor Professor Murray Burke for his supervision, encouragement and suggestions as well as the time he spent with me during the preparation of this thesis.

A special thank you is due to Dr. Tak Fung, Senior Statistical Consultant of Academic Computing Services, for his invaluable assistance.

The financial support of the University of Calgary through GAR/GAT is gratefully acknowledge.

Finally, I would like to express my thanks to Gisele Vezina–Facette for painstakingly and speedily typing the manuscript.

Dedicated to my parents

TABLE OF CONTENTS

Page

LIST OF TABLES

CHAPTER I
INTRODUCTION

Let X be a nonnegative random variable representing the time to failure (or death) of an article (or organism) with distribution function F(x) and probability density

$$f(x) = F'(x).$$

Then the conditional probability of failure during the next interval of duration x of an article (or organism) function at time t is

$$F(x|t) = \frac{F(t+x)-F(t)}{1-F(t)} .$$

Finally, we may obtain a conditional failure rate h(t) at time t:

$$h(t) = \lim_{\Delta x \to 0} \frac{1}{x} \frac{F(t+x)-F(x)}{1-F(t)}$$

$$= \frac{f(t)}{1-F(t)} . \qquad (1.1)$$

Alternate names for h(t) defined in (1.1) are hazard rate, mortality intensity, age–specific death rate, instantaneous death rate, and force of mortality. If the distribution of X is exponential with mean λ, then

$$h(x) = \frac{\lambda e^{-\lambda x}}{e^{-\lambda x}}$$

$$= \lambda$$

which is constant; for other distribution $h(x)$ varies with time.

There has been much interest in the above hazard rate. It is usefully utilized in reliability studies (Cox and Lewis, 1966), studies of mortality (Kimball, 1960) and in seismology (Gaisky, 1966; Udias and Rice, 1975). Many authors have dedicated tremendous effort to the estimation of the hazard rate. Watson and Leadbetter (1964a) suggested that the most obvious estimator combines the estimates of $f(x)$ and $F(x)$ based on a single random sample. Let us define, for $x_1,\ldots,x_n$ a simple random sample,

$$f_n(x) = \frac{1}{n}\sum_{i=1}^{n} \delta_n(x-X_i), \tag{1.2}$$

$$F_n(x) = \frac{1}{n}\,(\text{number of } X_i\text{'s} \leq x), \tag{1.3}$$

then

$$h_n^{(1)}(x) = \frac{f_n(x)}{1-F_n(x)}\,. \tag{1.4}$$

Estimator (1.2) has been proposed and discussed by Watson and Leadbetter (1961) and Parzen (1962). Here $\{\delta_n(x)\}$ is a δ–function sequence defined in section 1.1.

Let $X^{(1)} \leq X^{(2)} \leq \ldots \leq X^{(n)}$ be an ordered sample. If the data are assumed to come from an arbitrary distribution, the maximum–likelihood estimate of it in Grenander (1956) is a discrete distribution with probabilities $1/n$ at $X^{(i)}$. This gives a hazard rate of $1/(n-i)$ at $x = X^{(i)}$. To avoid the infinity at $i = n$ this estimate is changed to $1/(n - i + 1)$. To smooth linearly by using sequences of the smoothing function $\{\delta_n(x)\}$, Watson and

Leadbetter (1964a) derived

$$h_n^{(2)}(x) = \sum_{i=1}^{n} \frac{\delta_n(x-X^{(i)})}{n-i+1} \quad . \tag{1.5}$$

Rice and Rosenblatt (1976) further studied $h_n^{(1)}(x)$ and $h_n^{(2)}(x)$ and presented a similar estimator as follows:

$$h_n^{(3)}(x) = \sum_{i=1}^{n} \delta_n(x-X^{(i)})\log[1 + \frac{1}{n-i+1}].$$

Other nonparametric estimates of the hazard rate have been studied by Ahmad (1976), Ahmad and Lin (1977), Shaked (1978), and Miller and Singpurwalla (1978). However in this thesis I only provide a detailed and extensive analysis of the estimators $h_n^{(1)}(x)$, $h_n^{(2)}(x)$ and $h_n^{(3)}(x)$.

Appropriate definitions and preliminary results are described in the remainder of this chapter. Various convergence concepts and results also appear later.

1.1 SOME BASIC DEFINITIONS

Definition 1.1.1:

A sequence of functions $\{\delta_n(x)\}$ will be called a δ-function sequence if the following conditions hold:

(a) $\displaystyle\int_{-\infty}^{\infty} |\delta_n(x)|dx < A$, all n, some fixed A.

(b) $$\int_{-\infty}^{\infty} \delta_n(x)dx = 1, \text{ all } n$$

(1.1.1)

(c) $\delta_n(x) \longrightarrow 0$ uniformly in $|x| \geq \lambda$, for any fixed $\lambda > 0$.

(d) $$\int_{|x|\geq\lambda} |\delta_n(x)|dx \longrightarrow 0 \text{ as } n \longrightarrow \infty \text{ for any fixed } \lambda > 0.$$

A good example of such a δ–function sequence is

$$\delta_n(x) = w_n(x) = \frac{1}{b_n} w\left[\frac{x}{b_n}\right] \tag{1.1.2}$$

where w is given as a bounded, band limited, symmetric function of integral one and $b_n \downarrow 0$ with $nb_n \longrightarrow \infty$ as $n \longrightarrow \infty$. In this thesis, the above example (1.1.2) is often considered as we can derive more results by assuming

$$\delta_n(x) = \frac{1}{b_n} w \left[\frac{x}{b_n}\right].$$

For estimator $h_n^{(2)}(x)$, we further restrict the class of distributions. The introduction of such a class is for analytic convenience only.

<u>Definition 1.1.2</u>:

A class C_δ of distribution functions F(x) is such that for any fixed x_0, and any fixed $\lambda > 0$,

$$\frac{|\delta_n(x-x_0)|}{1-F(x)} \tag{1.1.3}$$

is, for all sufficiently large n, uniformly bounded in $|x-x_0| \geq \lambda$.

Here we consider G_λ as the upper bound of (1.1.3).

Later in Chapter 3, a weak approximation will come into use. Therefore, it is of interest to provide some basic settings for a weak approximation here.

A stochastic process is a collection $[X(t):t \in T]$ of random variables on a probability space $(\Omega,\mathcal{F},P)$. Usually T is thought of as representing time. In most cases, T is the set of integers and time is discrete, or else T is an interval of the line and time is continuous.

<u>Definition 1.1.3</u>:

A stochastic process $\{W(t;\omega) = W(t);\ 0 \leq t < \infty\}$, where $\omega \in \Omega$, and $\{\Omega,\mathcal{F},P\}$ is a probability space, is called a Wiener process if

(i) the process starts at 0, i.e, $W(0) = 0$,

(ii) the increment $W(t) - W(s)$ is normally distributed with mean 0 and variance $t - s$ for all $0 \leq s < t < \infty$,

(iii) $W(t)$ is an independent increment process, that is $W(t_2) - W(t_1)$, $W(t_4) - W(t_3),\ldots,W(t_{2n}) - W(t_{2n-1})$ are independent random variables for all $0 \leq t_1 < t_2 \leq t_3 < t_4 \leq\ldots\leq t_{2n-1} < t_{2n} < \infty$ $(n = 2,3,\ldots)$,

(iv) the sample path function $W(t;\omega)$ is continuous in t for all $\omega \in \Omega$.

Intuitively, (i) – (iii) imply that $E[W(t)] = 0$, $E[W^2(t)] = t$ and $E[W(s)W(t)] = s \wedge t$.

The proofs of the existence of a Wiener process and the continuity of its path can be found in Billingsley (1986).

Definition 1.1.4:

A random element $X(t)$, $0 \le t \le 1$, is Gaussian if all its finite–dimensional distributions are normal. On the other hand, the distribution of $X(t)$, $0 \le t \le 1$, is normal and the joint distribution of $X(t_1), X(t_2),\ldots,X(t_n)$, $0 \le t_1 < t_2 < \ldots < t_n \le 1$ $(n = 2,3,\ldots,)$ is multivariate normal. These finite–dimensional distributions are completely determined by the means $E[X(t)]$ and the covariance function $E[X(s)X(t)]$, $0 \le s,t \le 1$.

Definition 1.1.5:

A stochastic process $\{B(t);\ 0 \le t \le 1\}$ is called a Brownian bridge if

(i) the joint distribution of $B(t_1)$, $B(t_2),\ldots,B(t_n)$ $(0 \le t_1 < t_2 < \ldots < t_n \le 1;\ n = 1,2,\ldots)$ is Gaussian

(ii) $E[B(t)] = 0$ and $E[B(s)B(t)] = s \wedge t - st$, $0 \le s,t \le 1$,

(iii) the sample path function of $B(t;\omega)$ is continuous in t with probability one.

When $t = 0$ and 1, we have $E[B^2(t)] = 0$. Thus, $B(0) = B(1) = 0$ almost surely.

Csörgö and Rëvësz (1981) provided us the relationship between $B(t)$ and $W(t)$ as follows:

(i) Let $\{W(t); 0 \le t < \infty\}$ be a Wiener process. Then

$B(t) = W(t) - tW(1)$ $(0 \le t \le 1)$ is a Brownian bridge.

(1.1.4)

(ii) Let $B(t)$ be a Brownian bridge and define

$$W(t) = (t+1)B\left[\frac{t}{t+1}\right] \quad (0 \le t < \infty).$$

Then $W(t)$ is a Wiener process. (1.1.5)

The existence of such a Gaussian process, $B(t)$, follows immediately from (1.1.4).

With the idea of a Brownian bridge, we now go further to discuss strong approximation of the Empirical processes by such a Gaussian process. We first define

$$\{\beta_n(x); -\infty < x < \infty\} = \{\sqrt{n}(F_n(x) - F(x)); -\infty < x < \infty\}$$
$$(n = 1,2,\ldots) \qquad (1.1.6)$$

as the Empirical process. Then by the central limit theorem

$$\beta_n(x) \xrightarrow{\mathscr{D}} N(0,F(x)(1-F(x))) \qquad (1.1.7)$$

for each fixed x.

If F is continuous then $F^{-1}(y) = \inf \{x : F(x) = y\}$ and $F(F^{-1}(y)) = y$. As such, we can further investigate the process $\{\beta_n(x); -\infty < x < \infty\}$ by letting $U_i = F(X_i)$. Then the U_i are U(0,1) random variable provided F is continuous. Now let $E_n(y)$ be the empirical distribution of the sample $U_1,\ldots,U_n$ and the resulting uniform Empirical process is given by

$$\{\alpha_n(y); 0 \leq y \leq 1\} = \{\sqrt{n}(E_n(y) - y);\ 0 \leq y \leq 1\},\ n = 1,2,\ldots \tag{1.1.8}$$

Therefore, $\alpha_n(y) = \beta_n(F^{-1}(y))$, $0 \leq y \leq 1$. Further by (1.1.7) we have

$$\alpha_n(y) \xrightarrow{\mathscr{D}} N(0, y(1-y)) \tag{1.1.9}$$

for each fixed $y \in (0,1)$.

Finally, I display one of the best strong approximations of the Empirical process given by Komlós, Major and Tusnády (1975). For uniform empirical process, there exists a probability space on which one can define a sequence of Brownian bridges $\{B_n(y);\ 0 \leq y \leq 1\}$ such that

$$\sup_{0 \leq y \leq 1} |\alpha_n(y) - B_n(y)| \overset{\text{a.s.}}{=\!=} \mathcal{O}(n^{-\frac{1}{2}} \log n) \tag{1.1.10}$$

1.2 SOME PROBABILISTICAL RESULTS

<u>Theorem 1.2.1</u> (Chebyshev's Inequality):

Suppose that the random variable X has a distribution with mean μ and variance σ^2. Then for every $\epsilon > 0$,

$$P(|X-\mu| \geq \epsilon\sigma) \leq \frac{1}{\epsilon^2} .$$

Theorem 1.2.2 (The Glivenko–Cantelli lemma):

Suppose that $X_1,X_2,\ldots$ are independent and have a common–distribution F. Then

$$\sup_x |F_n(x) - F(x)| \xrightarrow{a.s.} 0 \tag{1.2.1}$$

as $n \longrightarrow \infty$. In words, F_n converges to F uniformly in x with probability 1.

Theorem 1.2.3 (Dominated Convergence Theorem):

If $|f_n| \leq g$ almost everywhere, where g is integrable and if $f_n \longrightarrow f$ almost everywhere, then f and the f_n are integrable and $\int f_n du \longrightarrow \int f\, du$.

Theorem 1.2.4 (Markov's Inequality):

Suppose X is a random variable, then for any $\epsilon > 0$ and $k \geq 0$,

$$P(|X| \geq \epsilon) \leq \frac{1}{\epsilon^k} E[|X|^k].$$

Theorem 1.2.5 (Hölder's Inequality):

Let X and Y be random variables and suppose that $\frac{1}{p} + \frac{1}{q} = 1$, $p > 1$, $q > 1$, then

$$E[|XY|] \leq \{E[|X|^p]\}^{\frac{1}{p}} \{E[|Y|^q]\}^{\frac{1}{q}} .$$

Theorem 1.2.6 (Normal Convergence Criterion):

If $X_{n,k}$ are independent summands, then, for every $\epsilon > 0$, the limiting distribution of $\sum_{k=1}^{n} X_{n,k}$ goes to $N(\alpha,\sigma^2)$ and $\max_k P[|X_{n,k}| \geq \epsilon] \longrightarrow 0$ if and only if, for every $\epsilon > 0$ and $\tau > 0$,

(i) $$\sum_{k=1}^{n} P[|X_{n,k}| \geq \epsilon] \longrightarrow 0$$

(ii) $$\sum_{k=1}^{n} \sigma_{n,k}^{2}(\tau) \longrightarrow \sigma^2, \quad \sum_{k=1}^{n} a_{n,k}(\tau) \longrightarrow \alpha.$$

Here $a_{n,k}(\tau) = \int_{|x|<\tau} x dF_{n,k}$, $\sigma_{n,k}^{2}(\tau) = \int_{|x|<\tau} x^2 dF_{n,k} - \left[\int_{|x|<\tau} x dF_{n,k}\right]$.

<u>Theorem 1.2.7</u> (C_r–inequality):

$E[|X+Y|^r] \leq C_r E[|X|^r] + C_r E[|Y|^r]$, where $C_r = 1$ or 2^{r-1} according as $r \leq 1$ or $r \geq 1$.

The proofs of theorem 1.2.1 – 1.2.7 can be easily found in many contexts of probability theory and hence will not be shown here. I have displayed these theorems here because they will be used later.

CHAPTER II
BIASEDNESS AND ITS ASYMPTOTIC RESULTS

In this chapter, I will examine the bias of the three aforementioned estimators. I first adapt the idea suggested by Rice and Rosenblatt (1976) of calculating the bias by assuming f is twice continuously differentiable, and then provide a more detailed analysis of the bias.

Watson and Leadbetter (1964b) has derived that $h_n^{(2)}$ is asymptotically unbiased. Considering the relation between the expected values of $h_n^{(2)}$ and $h_n^{(3)}$, I will also prove that $h_n^{(3)}$ is asymptotically unbiased.

In this chapter $O_p(\epsilon_n)$ has the usual meaning i.e. $A_n = O_p(\epsilon_n)$ means that A_n/ϵ_n are bounded random variables in probability for large n.

2.1 BIASEDNESS

Theorem 2.1.1 (Rice and Rosenblatt 1976):

Let $X_1,\ldots,X_n$ be independent random variables with common distribution F in C^1 (continuously differentiable) and density f.

Let $$a_n(x) = \frac{f_n(x)}{1-F(x)} \tag{2.1.1}$$

Then $$h_n^{(1)}(x) = a_n(x)\left[1 + O_p\left(\frac{1}{\sqrt{n}}\right)\right] \tag{2.1.2}$$

Choose $\delta_n(x) = w_n(x)$ and if $\int_{-\infty}^{\infty} w(x)x dx$ exists and f is twice continuously differentiable,

$$E[a_n(x)] = \frac{f(x)}{1-F(x)} + \frac{f''(x)}{1-F(x)} \int_{-\infty}^{\infty} w(v)v^2 dv \, \frac{b_n^2}{2} + o(b_n^2). \tag{2.1.3}$$

<u>Proof:</u>

$$h_n^{(1)}(x) = \frac{f_n(x)}{1-F(x)+[F(x)-F_n(x)]}$$

$$= f_n(x) \frac{\frac{1}{1-F(x)}}{1-\left[\frac{F_n(x)-F(x)}{1-F(x)}\right]}$$

$$= \frac{f_n(x)}{1-F(x)} \sum_{j=0}^{\infty} (-1)^j \left[\frac{F(x)-F_n(x)}{1-F(x)}\right]^j.$$

Since by (1.1.7),

$$\sup_x |F_n(x) - F(x)| = O_p\left(\frac{1}{\sqrt{n}}\right).$$

So

$$h_n^{(1)}(x) = \frac{f_n(x)}{1-F(x)} \left[1 + O_p\left(\frac{1}{\sqrt{n}}\right)\right]$$

$$= a_n(x)\left[1 + O_p\left(\frac{1}{\sqrt{n}}\right)\right].$$

This completes the proof of the first part.

Now,

$$E[a_n(x)] = \int_{-\infty}^{\infty} \delta_n(x-u) \frac{f(u)}{1-F(x)} \, du$$

$$= \int_{-\infty}^{\infty} \frac{1}{b_n} w \left[\frac{x-u}{b_n} \right] \frac{f(u)}{1-F(x)} \, du$$

$$= \int_{-\infty}^{\infty} w(v) \frac{f(x-b_n v)}{1-F(x)} \, dv.$$

Let

$$g(v) = f(x-b_n v),$$

then by the twice continuously differentiability of f and Taylor's Formula,

$$g(v) = g(0) + \frac{g'(0)}{1!} v + \frac{g''(0)}{2!} v^2 + R_2, \text{ where } R_2 \text{ is the error term}$$

$$= f(x) - b_n f'(x) v + \frac{b_n^2}{2} f''(x) v^2 + R_2.$$

As $R_2 \leq \frac{|b_n^3|}{3!} f'''(z)$, where z is between 0 and v,

$$E[a_n(x)] = \frac{f(x)}{1-F(x)} \int_{-\infty}^{\infty} w(v) dv - \frac{b_n^2 f'(x)}{1-F(x)} \int_{-\infty}^{\infty} w(v) v dv$$
$$+ \frac{b_n^2 f''(x)}{2[1-F(x)]} \int_{-\infty}^{\infty} w(v) v^2 dv + o(b_n^2).$$

Since w is a symmetric function of integral one,

$$E[a_n(x)] = \frac{f(x)}{1-F(x)} + \frac{f''(x)}{1-F(x)}\left[\int_{-\infty}^{\infty} w(v)v^2dv\right]\frac{b_n^2}{2} + o(b_n^2).$$

It follows from the above theorem, that the leading bias term of $h_n^{(1)}(x)$ is proportional to $f''(x)/(1-F(x))$. Noticing that

$$h'(x) = \frac{f'(x)}{1-F(x)} + \left[\frac{f(x)}{1-F(x)}\right]^2$$

and

$$h''(x) = \frac{f''(x)}{1-F(x)} + \frac{f'(x)f(x)}{[1-F(x)]^2} + \frac{2f'(x)f(x)}{[1-F(x)]^2} + 2\left[\frac{f(x)}{1-F(x)}\right]^3$$

$$= \frac{f''(x)}{1-F(x)} + \frac{3f'(x)h(x)}{[1-F(x)]^2} + 2(h(x))^3. \tag{2.1.4}$$

It is of interest to rewrite

$$\frac{f''(x)}{1-F(x)} = h''(x) - 3h(x)[h'(x) - (h(x))^2] - 2(h(x))^3$$

$$= h''(x) - 3h(x)\,h'(x) + (h(x))^3.$$

This rewritten expression shows for example that if $h'(x) = 0$ and $h''(x) > 0$ or if h is almost constant near x then the bias of $h_n^{(1)}(x)$ will be larger.

Following the idea of assuming f being twice continuously differentiable, we can also obtain the bias for $h_n^{(2)}$ and $h_n^{(3)}$ accordingly. To do this we need the following theorem.

<u>Theorem 2.1.2</u> (Rice and Rosenblatt (1976):

Let $X_1,\ldots,X_n$ be independent random variables with common distribution F in C^1 (continuously differentiable) and density f.

$$E[|h_n^{(2)}(x) - h_n^{(3)}(x)|] \leq \frac{k}{n}, \text{ some constant } k. \qquad (2.1.5)$$

Further choose $\delta_n(x) = w_n(x)$ and if $\int_{-\infty}^{\infty} w(x)x dx$ exists and f is twice continuously differentiable, then

$$\int_{-\infty}^{\infty} w_n(x-u) \frac{f(u)}{1-F(u)} du = \frac{f(x)}{1-F(x)} + \frac{b_n^2}{2} h''(x) \int_{-\infty}^{\infty} w(v)v^2 dv + o(b_n^2). \qquad (2.1.6)$$

<u>Proof</u>:

$$h_n^{(2)}(x) - h_n^{(3)}(x) = \sum_{i=1}^{n} \delta_n(x-X^{(i)})\left[\frac{1}{n-i+1} - \log(1 + \frac{1}{n-i+1}) \right].$$

Noticing that if $0 \leq x \leq 1$,

$$|x - \log(1 + x)| \leq |x - \sum_{j=1}^{\infty} \frac{x^j}{j} (-1)^{j+1}| \leq \frac{x^2}{2}.$$

Hence,

$$E[|h_n^{(2)}(x)-h_n^{(3)}(x)|] \leq \int_{-\infty}^{\infty} \sum_{i=1}^{n} \binom{n}{i-1} (F(u))^{i-1}(1-F(u))^{n-i} \frac{1}{n-i+1}\delta_n(x-u)f(u)du$$

$$\leq \int_{0<F(u)<1} \sum_{i=1}^{n} \begin{bmatrix} n \\ i-1 \end{bmatrix} F(u)^{i-1}(1-F(u))^{n-i} \frac{1}{n-i+1} \cdot$$

$$\delta_n(x-u)f(u)du + \int_{F(u)=1,0} \sum_{i=1}^{n} \begin{bmatrix} n \\ i-1 \end{bmatrix} F(u)^{i-1} \cdot$$

$$(1-F(u))^{n-i} \frac{1}{n-i+1} \delta_n(x-u)f(u)du$$

$$= \int_{0<F(u)<1} \sum_{i=1}^{n} \begin{bmatrix} n \\ i-1 \end{bmatrix} F(u)^{i-1} (1 - F(u))^{n-i} \frac{1}{n-i+1} \cdot$$
$$\delta_n(x-u)\ f(u)du$$

Now consider the expression $\sum_{i=1}^{n} \begin{bmatrix} n \\ i-1 \end{bmatrix} F(u)^{i-1}(1-F(u))^{n-i} \frac{1}{n-i+1}$. For $\alpha < 1$, the contribution from $\sum_{i<\alpha n}$ is $O\left[\frac{1}{n}\right]$.

For contribution of $\sum_{i\geq\alpha n}$, we consider F(u) as the probability of success in a binomial distribution of sample size n and variance nF(u) (1–F(u)). There is an $h > 0$ such that $F(u) \leq F(u + h) < \alpha < 1$. Let Y be the random variable with the above binomial distribution, then

$$\sum_{i\geq\alpha n} \begin{bmatrix} n \\ i-1 \end{bmatrix} F(u)^{i-1}(1-F(u))^{n-i} \frac{1}{(n-i+1)}$$

$$\leq \sum_{i\geq\alpha n} \begin{bmatrix} n \\ i-1 \end{bmatrix} F(u)^{i-1}(1-F(u))^{n-i}$$

$$\leq P[Y \geq \alpha n]$$

$$\leq P[|Y - nF(u)| \geq n(\alpha - F(u))].$$

By theorem (1.2.1) (Chebyshev's Inequality),

$$\sum_{i \geq \alpha n} \left[\begin{matrix} n \\ i-1 \end{matrix} \right] F(u)^{i-1}(1-F(u))^{n-i} \frac{1}{(n-i+1)} \leq \frac{F(u)(1-F(u))}{n(\alpha-F(u))^2}$$

$$\leq \frac{K}{n}, \text{ some constant } K.$$

Therefore,

$$E[|h_n^{(2)}(x) - h_n^{(3)}(x)|] \leq \frac{K}{n} \int_{-\infty}^{\infty} \delta_n(x-u)f(u)du.$$

By continuity f is bounded over $(-\infty,\infty)$. Let $f(u) \leq M$, $\forall u$. Then,

$$E[|h_n^{(2)}(x) - h_n^{(3)}(x)|] \leq \frac{KM}{n} \int_{-\infty}^{\infty} \delta_n(x-u)du$$

$$\leq \frac{KM}{n},$$

by (1.1.1(b)).

Hence,

$$E[|h_n^{(2)}(x) - h_n^{(3)}(x)|] \leq \frac{k}{n}, \text{ for some constant } k.$$

Now if f is twice continuously differentiable, then by using the method of Taylor's Formula as in the proof of theorem (2.1.1), we have

$$\int_{-\infty}^{\infty} w_n(x-u)\,\frac{f(u)}{1-F(u)}\,du$$

$$= \int_{-\infty}^{\infty} w(v)h(x-b_n v)dv$$

$$= \int_{-\infty}^{\infty} w(v)[h(x)-b_n h'(x)v + \frac{b_n^2}{2}\,h''(x)v^2+o(b_n^2)]dv$$

$$= h(x)\int_{-\infty}^{\infty} w(v)dv - b_n h'(x)\int_{-\infty}^{\infty} w(v)vdv$$

$$+ \frac{b_n^2}{2}\,h''(x)\int_{-\infty}^{\infty} w(v)v^2dv + o(b_n^2).$$

As w is a symmetric function of integral one,

$$\int_{-\infty}^{\infty} w_n(x-u)\frac{f(u)}{1-F(u)}\,du = h(x) + \frac{b_n^2}{2}\,h''(x)\int_{-\infty}^{\infty} w(v)v^2dv + o(b_n^2).$$

The proof of this theorem is completed.

Now choose $\delta_n(x) = w_n(x)$ and consider $h_n^{(2)}(x)$,

$$h_n^{(2)}(x) = \int_{-\infty}^{\infty} w_n(x-u) \frac{dF_n(u)}{1-F(u)+[F(u)-F_n(u)]}$$

$$= \int_{-\infty}^{\infty} w_n(x-u) \frac{1}{1-F(u)} \left[\sum_{j=0}^{\infty} (-1)^j \left[\frac{F(u)-F_n(u)}{1-F(u)} \right]^j \right] dF_n(u)$$

$$= \int_{-\infty}^{\infty} w_n(x-u) \frac{dF_n(u)}{1-F(u)} - \int_{-\infty}^{\infty} w_n(x-u) \frac{F(u)-F_n(u)}{[1-F(u)]^2} dF_n(u)$$

$$+ O_p\left[\frac{1}{n} \right] \qquad (2.1.7)$$

The mean of the first term on the right side of (2.1.7) is

$$\int_{-\infty}^{\infty} w_n(x-u) \frac{f(u)}{1-F(u)} du.$$

Thus by the preceding proved theorem (2.1.2), the leading bias term of $h_n^{(2)}(x)$ and $h_n^{(3)}(x)$ is proportional to $h''(x)$. From (2.1.4),

$$h''(x) = \frac{f''(x)}{1-F(x)} + \frac{3f'(x)h(x)}{[1-F(x)]^2} + 2(h(x))^3.$$

This expression shows, for example, that if $f'(x) = 0$ and $f''(x) > 0$ or if f is almost constant near x, the bias of $h_n^{(2)}(x)$ and $h_n^{(3)}(x)$ is greater.

2.2 ASYMPTOTIC UNBIASEDNESS

In this section, various results concerning the asymptotic behaviour of the bias of the three estimators will be studied here. Referring to $h_n^{(1)}(x)$, some results on the asymptotic behaviour of $f_n(x)$ and $F_n(x)$ have been given in Leadbetter and Watson (1961) and Parzen (1962) and they have certain implications for asymptotic unbiasedness of $h_n^{(1)}(x)$.

All theorems and lemmas except theorem (2.2.3) in this section are extracted from Watson and Leadbetter (1964b).

<u>Lemma 2.2.1</u>:

If g(x) is continuous at x = 0 and g(x) is integrable, and if $\{\delta_n(x)\}$ is a δ–function sequence, then $g(x)\delta_n(x)$ is integrable and $\int_{-\infty}^{\infty} g(x)\delta_n(x)dx \longrightarrow g(0)$ as $n \longrightarrow \infty$.

<u>Proof</u>:

By continuity g(x) is bounded in some interval $(-\lambda,\lambda)$ with $\lambda > 0$. Suppose $g(x) \le M$, for some fixed M whenever $x \in (-\lambda,\lambda)$, then

$$\int_{|x|<\lambda} g(x)\delta_n(x)dx \le M \int_{|x|<\lambda} \delta_n(x)dx.$$

Hence $g(x)\delta_n(x)$ is integrable over $(-\lambda,\lambda)$ by (1.1.1(a)). That $g(x)\delta_n(x)$ is also integrable over the region $|x| \ge \lambda$ follows from (1.1.1(c)) and integrability of g. Now let A be chosen as in (1.1.1(a)). Then for $\epsilon > 0$, λ may be chosen such that $|g(x) - g(0)| < \epsilon/A$ if $|x| < \lambda$. Thus,

$$\left|\int_{-\infty}^{\infty} g(x)\delta_n(x)dx - g(0)\right| = \left|\int_{-\infty}^{\infty} g(x)\delta_n(x)dx - \int_{-\infty}^{\infty} g(0)\delta_n(x)dx\right|$$

by (1.1.1(b))

$$= \left| \int_{-\infty}^{\infty} [g(x) - g(0)]\delta_n(x)dx \right|$$

$$\leq \int_{-\infty}^{\infty} |g(x) - g(0)| \; |\delta_n(x)|dx$$

$$\leq \int_{|x|<\lambda} |g(x) - g(0)| \; |\delta_n(x)|dx + \int_{|x|\geq\lambda} |g(x) - g(0)| \; |\delta_n(x)|dx$$

$$\leq \int_{|x|<\lambda} |g(x) - g(0)| \; |\delta_n(x)|dx + \int_{|x|\geq\lambda} [|g(x)| + |g(0)|] \; |\delta_n(x)|dx$$

$$= \int_{|x|<\lambda} |g(x) - g(0)| \; |\delta_n(x)|d_x + \int_{|x|\geq\lambda} |g(x)| \cdot |\delta_n(x)|dx + |g(0)| \int_{|x|\geq\lambda} |\delta_n(x)|dx. \qquad (2.2.1)$$

The first term of (2.2.1)

$$\int_{|x|<\lambda} |g(x) - g(0)| \; |\delta_n(x)|dx < \int_{|x|<\lambda} \frac{\epsilon}{A} |\delta_n(x)|dx$$

$$= \frac{\epsilon}{A} \int_{|x|<\lambda} |\delta_n(x)|dx$$

$$< \frac{\epsilon}{A} \times A \qquad \text{by } (1.1.1(a))$$

$$= \epsilon.$$

We let $K_n = \sup_{|x|\geq\lambda} |\delta_n(x)|$, then the second term of (2.2.1)

$$\int_{|x|\geq\lambda} |g(x)|\ |\delta_n(x)|\,dx \leq K_n \int_{|x|\geq\lambda} |g(x)|\,dx.$$

By (1.1.1(b)), $K_n \to 0$ as $n \to \infty$. Follows by the integrability of g,

$$\int_{|x|\geq\lambda} |g(x)|\ |\delta_n(x)|\,dx \to 0 \text{ as } n \to \infty.$$

By (1.1.1(d)),

$$|g(0)| \int_{|x|\geq\lambda} |\delta_n(x)|\,dx \to 0 \text{ as } n \to \infty.$$

Since ϵ is arbitrary,

$$\int_{-\infty}^{\infty} g(x)\ \delta_n(x) \to g(0) \text{ as } n \to \infty.$$

This completes the proof of this lemma.

By this important lemma, we have the following result.

Theorem 2.2.1:

Let $\{\delta_n(x)\}$ be a δ–function sequence and f(x) a probability density which is continuous at the point x_0. Let $f_n(x_0) = \frac{1}{n} \sum_{i=1}^{n} \delta_n(x_0 - X_i)$ where $X_1,\ldots,X_n$

form an independent sample from the distribution. Then

$$E[f_n(x_0)] \longrightarrow f(x_0) \text{ as } n \longrightarrow \infty.$$

Proof:

Since

$$E[f_n(x_0)] = \int_{-\infty}^{\infty} \delta_n(x_0 - u)f(u)du$$

$$= \int_{-\infty}^{\infty} \delta_n(v)f(x_0 - v)dv.$$

By Lemma (2.2.1), $E[f_n(x_0)] \longrightarrow f(x_0)$ as $n \longrightarrow \infty$.

The sample distribution function $F_n(x)$ is essentially a binomially distributed random variable with mean $F(x)$. Therefore,

$$E[F_n(x)] = F(x) \qquad (2.2.2)$$

and

$$\text{var}[F_n(x)] = \frac{1}{n}F(x)[1-F(x)]. \qquad (2.2.3)$$

So $F_n(x)$ is an unbiased estimator of $F(x)$. By the above results, we can obtain the asymptotic unbiasedness of $h_n^{(1)}(x)$. But since the asymptotic unbiasedness of $h_n^{(1)}(x)$ follows immediately from the asymptotic normality of $h_n^{(1)}(x)$, we discuss it later in Chapter 3.

We now turn our attention to $h_n^{(2)}(x)$.

$$E[h_n^{(2)}(x)] = E\left[\sum_{i=1}^{n} \frac{1}{n-i+1}\, \delta_n(x-X^{(i)})\right]$$

$$= \int_{-\infty}^{\infty} \left\{\sum_{i=1}^{n} \begin{bmatrix} n \\ i-1 \end{bmatrix} F(u)^{i-1}[1-F(u)]^{n-i}\right\} \delta_n(x-u)f(u)du$$

$$= \int_{-\infty}^{\infty} \left\{\sum_{i=1}^{n+1} \begin{bmatrix} n \\ i-1 \end{bmatrix} F(u)^{i-1}[1-F(u)]^{n-i+1}\right\} \left\{\frac{1}{1-F(u)} - F(u)^{n}\right\}. \\ \delta_n(x-u)f(u)du$$

$$= \int_{-\infty}^{\infty} \left\{\sum_{r=0}^{n} \begin{bmatrix} n \\ r \end{bmatrix} F(u)^{r}\, [1-F(u)]^{n-r}\right\} \left\{\frac{1}{1-F(u)} - F(u)^{n}\right\}. \\ \delta_n(x-u)f(u)du.$$

Considering F(u) as the mean of a binomial distribution, then

$$\sum_{r=0}^{n} \begin{bmatrix} n \\ r \end{bmatrix} F(u)^{r}\, [1-F(u)]^{n-r} = 1.$$

That is

$$E[h_n^{(2)}(x)] = \int_{-\infty}^{\infty} \left\{\frac{1}{1-F(u)} - F(u)^{n}\right\} \delta_n(x-u)f(u)du$$

$$= \int_{-\infty}^{\infty} \delta_n(x-u)h(u)du - \int_{-\infty}^{\infty} \delta_n(x-u)F(u)^n f(u)du \tag{2.2.4}$$

Lemma 2.2.2:

If $\{\delta_n\}$ is a δ–function sequence and if F(x) is a distribution function in the class C_δ, then, provided the associated hazard function h(x) is continuous at x_0, we have

$$\int_{-\infty}^{\infty} \delta_n(x_0-u)h(u)du \longrightarrow h(x_0) \text{ as } n \longrightarrow \infty.$$

Proof:

For given $\epsilon > 0$, we choose $\lambda > 0$ such that $|h(u) - h(x_0)| < \frac{\epsilon}{A}$ if $|u - x_0| < \lambda$. Then,

$$\left|\int_{-\infty}^{\infty} \delta_n(u - x_0)h(u)du - h(x_0)\right|$$

$$= \left|\int_{-\infty}^{\infty} \delta_n(u - x_0)h(u)du - \int_{-\infty}^{\infty} \delta_n(u - x_0)h(x_0)du\right| \qquad \text{by (1.1.1(b))}$$

$$= \left|\int_{-\infty}^{\infty} \delta_n(u - x_0)[h(u) - h(x_0)]du\right|$$

$$= \left|\int_{|u-x_0|<\lambda} \delta_n(u-x_0)[h(u)-h(x_0)]du + \int_{|u-x_0|\geq\lambda} \delta_n(u-x_0)[h(u)-h(x_0)]du\right|$$

$$\leq \left|\int_{|u-x_0|<\lambda} \delta_n(u-x_0)[h(u)-h(x_0)]du\right| + \left|\int_{|u-x_0|\geq\lambda} \delta_n(u-x_0)[h(u)-h(x_0)]du\right|$$

$$\leq \int_{|u-x_0|<\lambda} |\delta_n(u-x_0)|\,|h(u)-h(x_0)|\,du + \int_{|u-x_0|\geq\lambda} |\delta_n(u-x_0)|\,|h(u)|\,du$$
$$+ \int_{|u-x_0|\geq\lambda} |\delta_n(u-x_0)| \quad |h(x_0)|\,du$$

$$< \frac{\epsilon}{A} \int_{|u-x_0|<\lambda} |\delta_n(u-x_0)|\,du + \int_{|u-x_0|\geq\lambda} |\delta_n(u-x_0)| \quad |h(u)|\,du$$
$$+ \quad |h(x_0)| \int_{|u-x_0|\geq\lambda} |\delta_n(u-x_0)|\,du$$

$$< \epsilon + \int_{|u-x_0|\geq\lambda} |\delta_n(u-x_0)| \quad |h(u)|\,du + |h(x_0)| \int_{|u-x_0|\geq\lambda} |\delta_n(u-x_0)|\,du \tag{2.2.5}$$

By definition (1.1.2), the integrand of the second term of (2.2.5) is dominated by $G_\lambda f(u)$. From (1.1.1(c)), this integrand tends to zero for all $|u-x_0| \geq \lambda$. It follows from the dominated convergence theorem (1.2.3) that the second term tends to zero. Furthermore, the third term of (2.2.5) also tends to zero simply by (1.1.1(d)). Since ϵ is arbitrary, the result follows.

<u>Theorem 2.2.2</u>:

Let $\{\delta_n\}$ be a δ–function sequence and $F(x)$ a distribution function in the class C_δ. If the hazard function $h(x)$ is continuous at x_0, and if $F(x_0) < 1$, then $h_n^{(2)}(x_0)$ is an asymptotically unbiased estimator of $h(x_0)$.

<u>Proof</u>:

By (2.2.4) and lemma (2.2.2), it is sufficient to show that

$$\int \delta_n(x-u)F(u)^n f(u)du \longrightarrow 0 \text{ as } n \longrightarrow \infty.$$

If $\lambda > 0$ is chosen so that $F(x_0+\lambda) < 1$ and further $h(u)$ is bounded in $|u-x_0| < \lambda$, then

$$\int_{|x_0-u|\geq\lambda} \delta_n(x_0-u)F(u)^n f(u)du$$

$$= \int_{|x_0-u|\geq\lambda} \delta_n(x_0-u)h(u)[1-F(u)]\ F(u)^n du$$

$$\leq G_\lambda \int_{|x_0-u|\geq\lambda} [F(u)^n - F(u)^{n+1}]\ f(u)du$$

$$\leq G_\lambda \int_0^1 [F(u)^n - F(u)^{n+1}]dF(u)$$

$$\leq G_\lambda \left[\frac{1}{n+1} - \frac{1}{n+2} \right]$$

$$\leq G_\lambda \left[\frac{1}{(n+1)(n+2)} \right] \longrightarrow 0 \text{ as } n \longrightarrow \infty.$$

Further since $h(u)$ is bounded in $|x_0 - u| < \lambda$,

$$\int_{|x_0-u|<\lambda} \delta_n(x_0-u)F(u)^n[1-F(u)]h(u)du$$

$$\leq \int_{|x_0-u|<\lambda} \delta_n(x_0-u)F(u)^n K_1 du, \qquad \text{some constant } K_1$$

$$\leq K_1 \int_{|x_0-u|<\lambda} \delta_n(x_0-u)F(x_0+\lambda)^n du$$

$$= K_1 F(x_0+\lambda)^n \int_{|x_0-u|<\lambda} \delta_n(x_0-u)du$$

$$\leq K_1 K_2 F(x_0+\lambda)^n \qquad \text{by} \quad (1.1.1(a))$$

which goes to zero as n goes to infinity. Combining the above two results, we have

$$\int_{-\infty}^{\infty} \delta_n(x-u)F(u)^n f(u)du \longrightarrow 0 \text{ as } n \longrightarrow \infty,$$

and the theorem follows.

To show that $h_n^{(3)}(x)$ is also asymptotically unbiased, we should recall (2.1.5) from theorem (2.1.2) wherein

$$E[|h_n^{(2)}(x) - h_n^{(3)}(x)|] \leq \frac{k}{n} \longrightarrow 0 \text{ as } n \longrightarrow \infty.$$

Hence,

$$E[|h_n^{(2)}(x) - h_n^{(3)}(x)|] \longrightarrow 0 \text{ as } n \longrightarrow \infty$$

implying that $h_n^{(3)}(x)$ is also asymptotically unbiased by theorem (2.2.2). But we should note that a further condition (continuously differentiability of F) is required here in accord with theorem (2.1.2). So we have the following theorem.

Theorem 2.2.3:

Let $X_1,\ldots,X_n$ be independent random variables with common distribution F in C^1 and in the class C_δ. Let $\{\delta_n\}$ be a δ–function sequence, then $h_n^{(3)}(x)$ is an asymptotically unbiased estimator of $h(x)$.

CHAPTER III
ASYMPTOTIC NORMALITY AND DEVIATION

Having examined the bias and its asymptotic behaviour of the three estimators in the previous chapter, I will now obtain other asymptotical results. In section 3.1, we show that if $\delta_n(x) = w_n(x)$ and the weight function w of (1.1.2) has finite support then all three estimators are asymptotically equivalent. In section 3.2, we then obtain the asymptotic normality of $h_n^{(1)}(x)$ by extracting some theorems and lemmas from Watson and Leadbetter (1964b). However, in the final section, we switch our concentration to the study of the global deviations of $h_n^{(2)}(x)$ and $h_n^{(3)}(x)$ via some convergence theorems.

3.1 ASYMPTOTIC EQUIVALENCE

<u>Theorem 3.1.1</u>:

Let $X_1,\ldots,X_n$ be independent random variables with common distribution F in C^1 (continuously differentiable) and density f. Choose $\delta_n(x) = w_n(x)$ and further if w has finite support, say vanishing outside $[-A,A]$, then

$$|h_n^{(2)}(x) - h_n^{(1)}(x)| = O_p(\max(n^{-\frac{1}{2}}, b_n)).$$

<u>Proof</u>:

$|h_n^{(2)}(x) - h_n^{(1)}(x)|$

$$= \left| \int_{-A}^{A} \frac{1}{b_n} w\left[\frac{x-u}{b_n} \right] \frac{1}{1-F_n(u)} \, dF_n(u) - \int_{-A}^{A} \frac{1}{b_n} w\left[\frac{x-u}{b_n} \right] \frac{1}{1-F_n(x)} \, dF_n(u) \right|$$

$$= \left| \frac{1}{b_n} \int_{-A}^{A} w\left[\frac{x-u}{b_n} \right] \left[\frac{1}{1-F_n(u)} - \frac{1}{1-F_n(x)} \right] dF_n(u) \right|$$

$$= \left| \frac{1}{b_n} \int_{-A}^{A} w\left[\frac{x-u}{b_n} \right] \left\{ \frac{F_n(u) - F_n(x)}{[1-F_n(u)][1-F_n(x)]} \right\} dF_n(u) \right|$$

$$= \left| \frac{1}{b_n} \int_{-A}^{A} w(y) \left\{ \frac{F_n(x-b_n y) - F_n(x)}{[1-F_n(x-b_n y)][1-F_n(x)]} \right\} b_n dF_n(x-b_n y) \right|, \text{ by putting } y = \frac{x-u}{b_n}$$

$$\le \int_{-A}^{A} |w(y)| \frac{|[F_n(x-b_n y) - F(x-b_n y)| + |F_n(x) - F(x)| + |F(x-b_n y) - F(x)|}{|1 - F_n(x-b_n y)| \, |1-F_n(x)|} \cdot dF_n(x-b_n y) \qquad (3.1.1)$$

By (1.1.7),

$$\sup_x |F_n(x-b_n y) - F(x-b_n y)| = O_p\left(\frac{1}{\sqrt{n}}\right)$$

and

$$\sup_x |F_n(x) - F(x)| = O_p\left(\frac{1}{\sqrt{n}}\right).$$

Also by mean value theorem,

$$|F(x - b_n y) - F(x)| \le |b_n y\, f(\xi)|, \text{ where } f(\xi) = \max_{x - b_n|y| \le t \le x + b_n|y|} f(t)$$

$$\le b_n A |f(\xi)|.$$

Hence

$$|F(x - b_n y) - F(x)| = O(b_n).$$

Thus,

$$|h_n^{(2)}(x) - h_n^{(1)}(x)| = O_p(\max(n^{-\frac{1}{2}}, b_n)).$$

<u>Theorem 3.1.2</u> (Rice and Rosenblatt 1976):

Let $X_1, \ldots, X_n$ be independent random variables with common distribution F in C^1 (continuously differentiable) and density f, then

$$|h_n^{(2)}(x) - h_n^{(3)}(x)| = O_p\left[\frac{1}{n}\right].$$

<u>Proof</u>:

The proof follows immediately from (2.1.5) in theorem (2.1.2) where

$$E[|h_n^{(2)}(x) - h_n^{(3)}(x)|] \le \frac{k}{n}, \text{ some constant } k$$

then by Markov's inequality (1.2.4),

$$P\ \{|h_n^{(2)}(x) - h_n^{(3)}(x)| \ge \epsilon\} \le \frac{1}{\epsilon} E[|h_n^{(2)}(x) - h_n^{(3)}(x)|]$$

$$\le \frac{k}{n\epsilon} \longrightarrow 0 \text{ as } n \longrightarrow \infty.$$

By the above two significant theorems, we can conclude that all three estimators are asymptotically equivalent if $\delta_n(x) = w_n(x)$ and w has finite support.

3.2 ASYMPTOTIC NORMALITY

The asymptotic normal behaviour of $h_n^{(1)}(x)$ will be given here through a series of lemmas and theorems without the assumption that $\delta_n = w_n(x)$.

Lemma 3.2.1:

If $\{\delta_n\}$ is a δ-function sequence, and for $p \geq 2$, $\alpha_n = \alpha_n(p) = \int_{-\infty}^{\infty} |\delta_n(x)|^p dx < \infty$, then $\alpha_n \longrightarrow \infty$ as $n \longrightarrow \infty$.

Proof:

It follows from (1.1.1(b) & (d)) that if $\lambda > 0$, $\int_{-\lambda}^{\lambda} \delta_n(x)dx \longrightarrow 1$ as $n \longrightarrow \infty$. Now

$$\left| \int_{-\lambda}^{\lambda} \delta_n(x)dx \right| \leq \int_{-\lambda}^{\lambda} |\delta_n(x)|dx.$$

By Hölders inequality (1.2.5) where $q^{-1} = 1 - p^{-1}$,

$$\left| \int_{-\lambda}^{\lambda} \delta_n(x)dx \right| \leq (2\lambda)^{\frac{1}{q}} \left\{ \int_{-\lambda}^{\lambda} |\delta_n(x)|^p dx \right\}^{\frac{1}{p}}.$$

Therefore,

$$(2\lambda)^{\frac{1}{q}} \liminf_{n \to \infty} \left\{\int_{-\infty}^{\infty} |\delta_n(x)|^p dx\right\}^{\frac{1}{p}} = \liminf_{n \to \infty} (2\lambda)^{\frac{1}{q}} \left\{\int_{-\infty}^{\infty} |\delta_n(x)|^p dx\right\}^{\frac{1}{p}}$$

$$\geq \liminf_{n \to \infty} \left|\int_{-\lambda}^{\lambda} \delta_n(x) dx\right|$$

$$= 1.$$

Thus,

$$\liminf_{n \to \infty} \left\{\int_{-\infty}^{\infty} |\delta_n(x)|^p dx\right\}^{\frac{1}{p}} \geq \frac{1}{(2\lambda)^{\frac{1}{q}}}.$$

This is true for all $\lambda > 0$, hence $\alpha_n = \int_{-\infty}^{\infty} |\delta_n(x)|^p dx \to \infty$ as $n \to \infty$.

<u>Lemma 3.2.2</u>:

Let $\{\delta_n(x)\}$ be a δ–function sequence with $\alpha_n = \alpha_n(p) = \int_{-\infty}^{\infty} |\delta_n(x)|^p dx$ $< \infty$. Then $\overset{*}{\delta}_n(x) = \frac{|\delta_n(x)|^p}{\alpha_n}$ is also a δ–function sequence, $(p \geq 2)$.

<u>Proof</u>:

$\overset{*}{\delta}_n$ satisfies condition (1.1.1(a) & (b)) as

$$\int_{-\infty}^{\infty} |\overset{*}{\delta}_n(x)|\,dx = \int_{-\infty}^{\infty} \overset{*}{\delta}_n(x)dx$$

$$= \alpha_n^{-1}\int_{-\infty}^{\infty} |\delta_n(x)|^p dx$$

$$= 1.$$

$\overset{*}{\delta}_n$ satisfies (1.1.1(c)) since $\delta_n(x) \longrightarrow 0$ uniformly in $|x| \geq \lambda > 0$ and $\alpha_n \longrightarrow \infty$ by lemma (3.2.1). Now let $K_n = \sup_{|x|\geq\lambda} |\delta_n(x)|$, then

$$\int_{|x|\geq\lambda} |\overset{*}{\delta}_n(x)|\,dx \leq \left[\frac{K_n{}^{p-1}}{\alpha_n}\right] \int_{|x|\geq\lambda} |\delta_n(x)|\,dx.$$

Since $K_n \longrightarrow 0$, $\alpha_n \longrightarrow \infty$ and $\int_{|x|\geq\lambda} |\delta_n(x)|\,dx \longrightarrow 0$ as $n \longrightarrow \infty$,

$\int_{|x|\geq\lambda} |\overset{*}{\delta}_n(x)|\,dx \longrightarrow 0$ as $n \longrightarrow \infty$ i.e. (1.1.1(d)) is also satisfied by $\overset{*}{\delta}_n$. The lemma follows.

<u>Theorem 3.2.1</u>:

Let $\{\delta(x)\}$ be a δ-function sequence with $\alpha_n = \int_{-\infty}^{\infty} \delta_n^2(x)dx < \infty$ and $\alpha_n = o(n)$. Let x_0 be a continuity point of the density f. Then $(n/\alpha_n)\text{var}\{f_n(x_0)\} \longrightarrow f(x_0)$ as $n \longrightarrow \infty$.

Proof:

$$\mathrm{var}[f_n(x_0)] = \frac{1}{n} E[\delta_n^2(x_0-X_1)] - \frac{1}{n}\{E[\delta_n(x_0-X_1)]\}^2$$

$$\Rightarrow \left[\frac{n}{\alpha_n}\right] \mathrm{var}[f_n(x_0)] = \alpha_n^{-1}\int_{-\infty}^{\infty} \delta_n^2(x_0-u)f(u)du - \alpha_n^{-1}\left\{\left[\int_{-\infty}^{\infty} \delta_n(x_0-u)f(u)du\right\}\right]^2.$$

By lemma (3.2.2), $\overset{*}{\delta_n} = \delta_n^2/\alpha_n$ is also a δ-function sequence, then

$$\alpha_n^{-1}\int_{-\infty}^{\infty} \delta_n^2(x_0-u)f(u)du = \int_{-\infty}^{\infty} \overset{*}{\delta_n}(x_0-u)f(u)du \longrightarrow f(x_0) \text{ as } n \longrightarrow \infty$$

by lemma (2.2.1),

and $\alpha_n^{-1}\left\{\left[\int_{-\infty}^{\infty} \delta_n(x_0-u)f(u)du\right]\right\}^2 \longrightarrow 0$ by lemmas (2.2.1) and (3.2.1).

Hence, the theorem.

Theorem 3.2.2:

Let $\{\delta_n(x)\}$ be a δ-function sequence such that

$$\alpha_n = \int_{-\infty}^{\infty} \delta_n^2(x)dx < \infty,\ \gamma_n = \int_{-\infty}^{\infty} |\delta_n(x)|^{2+\eta}dx < \infty \text{ for some } \eta > 0,$$

and such that $\dfrac{\gamma_n}{n^{\eta/2}\alpha_n^{1+\eta/2}} \longrightarrow 0$ as $n \longrightarrow \infty$. Let x be a continuity point of the probability density f. Then $n^{\frac{1}{2}}\{f_n(x) - E[f_n(x)]\}/[\alpha_n f(x)]^{\frac{1}{2}}$ has the standard normal limiting distribution.

Proof:

Let $Z_n = \dfrac{\{f_n(x) - E[f_n(x)]\}}{\{var[f_n(x)]\}^{\frac{1}{2}}}$, then by theorem (3.2.1) it is sufficient to show that Z_n has the standard normal limiting distribution. But we can rewrite

$$Z_n = \sum_{i=1}^{n} X_{n,i}$$

where

$$\mu_n = E[\delta_n(x-X_1)],\ \sigma_n^2 = var[\delta_n(x-X_1)],$$

$$X_{n,i} = [\delta_n(x-X_i) - \mu_n]/(n^{\frac{1}{2}}\sigma_n).$$

Hence,

$$\sum_{i=1}^{n} E[X_{n,i}] = 0 \text{ and } \sum_{i=1}^{n} var[X_{n,i}] = 1.$$

By normal convergence criterion (1.2.6), it is sufficient to show that $\sum_{i=1}^{n} P[|X_{n,i}| \geq \epsilon] \longrightarrow 0$ as $n \longrightarrow \infty$. If $F_{n,i}$ is the distribution function of $X_{n,i}$, then by Markov's inequality (1.2.4)

$$\sum_{i=1}^{n} P[|X_{n,i}| \geq \epsilon] \leq \frac{n}{\epsilon^{2+\eta}} E[|X_{n,i}|^{2+\eta}]$$

$$= \frac{n}{\epsilon^{2+\eta}} E\left\{\left| \frac{\delta_n(x-X_1)-E[\delta_n(x-X_1)]}{[n\ var\ \delta_n(x-X_1)]^{\frac{1}{2}}}\right|^{2+\eta}\right\}.$$

Hence, we only need to show that

$$n\ E\left\{\left|\frac{\delta_n(x-X_1)-E[\delta_n(x-X_1)]}{[n\ \mathrm{var}\ \delta_n(x-X_1)]^{\frac{1}{2}}}\right|^{2+\eta}\right\} \longrightarrow 0 \text{ as } n \longrightarrow \infty.$$

Following from the C_r–inequality (1.2.7) given in Loève (1960), this expression is dominated by

$$2^{1+\eta}\left\{\frac{E|\delta_n(x-X_1)|^{2+\eta}}{n^{\eta/2}[\mathrm{var}(\delta_n(x-X_1))]^{1+\eta/2}}+\frac{|E[\delta_n(x-X_1)]|^{2+\eta}}{n^{\eta/2}[\mathrm{var}(\delta_n(x-X_1))]^{1+\eta/2}}\right\}$$

$$= 2^{1+\eta}[P_{1,n} + P_{2,n}] \text{ say.}$$

By theorems (2.2.1) and (3.2.2),

$$P_{2,n} \longrightarrow \frac{|f(x)|^{2+\eta}}{n^{\eta/2}[\alpha_n f(x)]^{1+\eta/2}} \longrightarrow 0 \text{ as } n \longrightarrow \infty \text{ by lemma (3.2.1).}$$

Also,

$$P_{1,n} \longrightarrow \frac{E|\delta_n(x-X_1)|^{2+\eta}}{n^{\eta/2}[\alpha_n f(x)]^{1+\eta/2}} \quad \text{as } n \longrightarrow \infty.$$

By lemma (3.2.2), $\overset{*}{\delta}_n(x) = \dfrac{|\delta_n(x-X_1)|^{2+\eta}}{\gamma_n}$ is also a δ–function sequence. Then

$$P_{1,n} \longrightarrow \frac{E[\overset{*}{\delta}_n(x-X_1)]\gamma_n}{n^{\eta/2}\alpha_n^{1+\eta/2}f(x)^{1+\eta/2}} \longrightarrow \frac{f(x)\gamma_n}{n^{\eta/2}\alpha_n^{1+\eta/2}f(x)^{1+\eta/2}} \text{ as } n \longrightarrow \infty$$

by theorem (2.2.1).

Since by hypothesis $(\gamma_n/n^{\eta/2}\alpha_n^{1+\eta/2} \longrightarrow 0$ as $n \longrightarrow \infty)$,

$$P_{1,n} \longrightarrow 0 \text{ as } n \longrightarrow \infty.$$

The theorem follows.

<u>Theorem 3.2.3</u>:

Let $f(x)$ be the probability density of a non–negative random variable. Under the conditions of theorem (3.2.2) and $\left[\frac{n}{\alpha_n}\right]^{\frac{1}{2}}\{E[f_n(x)] - f(x)\} \longrightarrow 0$ as $n \longrightarrow \infty$, then the random variable Y_n defined by

$$Y_n = [1 - F(x)]\ [n/\alpha_n f(x)]^{\frac{1}{2}}\ [h_n^{(1)}(x) - h(x)]$$

has the standard normal limiting distribution.

<u>Proof</u>:

$$\text{Let } W_n = [1-F(x)] \left[\frac{n}{\alpha_n f(x)}\right]^{\frac{1}{2}} \left\{ h_n^{(1)}(x) - \frac{E[f_n(x)]}{1-F_n(x)} \right\}$$

$$= \left[\frac{n}{\alpha_n f(x)}\right]^{\frac{1}{2}} \left[\frac{1-F(x)}{1-F_n(x)}\right] \left\{f_n(x) - E[f_n(x)]\right\}$$

$$= \left[\frac{1-F(x)}{1-F_n(x)}\right] Z_n.$$

By theorem (1.2,2), $\frac{1-F(x)}{1-F_n(x)} \longrightarrow 1$ in probability. Following from the limiting standard normal distribution of Z_n as in theorem (3.2.2), W_n converges to a standard normal random variable. But

$$Y_n = W_n + \left[\frac{n}{\alpha_n f(x)}\right]^{\frac{1}{2}} [1-F(x)] \left\{\frac{E[f_n(x)]}{1-F_n(x)}\right\} - \left[\frac{n}{\alpha_n f(x)}\right]^{\frac{1}{2}} f(x)$$

$$= W_n + \left[\frac{n}{\alpha_n f(x)}\right]^{\frac{1}{2}} [1-F(x)] \left\{\frac{E[f_n(x)]}{1-F_n(x)}\right\}$$

$$+\left[\frac{n}{\alpha_n f(x)}\right]^{\frac{1}{2}} \{E[f_n(x)] - f(x)\} - \left[\frac{n}{\alpha_n f(x)}\right]^{\frac{1}{2}} E[f_n(x)]$$

$$= W_n + \left[\frac{n}{\alpha_n f(x)}\right]^{\frac{1}{2}} \left\{\frac{E[f_n(x)]-F(x)E[f_n(x)]-E[f_n(x)]+F_n(x)E[f_n(x)]}{1-F_n(x)}\right\}$$

$$+\left[\frac{n}{\alpha_n f(x)}\right]^{\frac{1}{2}} \{E[f_n(x)] - f(x)\}$$

$$= W_n + \left[\frac{n}{\alpha_n f(x)}\right]^{\frac{1}{2}} \left[\frac{F_n(x)-F(x)}{1-F_n(x)}\right] E[f_n(x)] + \left[\frac{n}{\alpha_n f(x)}\right]^{\frac{1}{2}} \{E[f_n(x)] - f(x)\} \tag{3.2.1}$$

By (1.1.7) and lemma (3.2.1),

$\left[\frac{n}{\alpha_n}\right]^{\frac{1}{2}} [F_n(x) - F(x)] \to 0$ as $n \to \infty$. Hence the second term of (3.2.1) tends to zero by theorem (2.2.1). The last term of (3.2.1) also tends to zero as n goes to infinity by the hypothesis. So Y_n has the same limiting distribution as W_n i.e. Y_n has the standard normal limiting distribution. This completes the proof.

Under the conditions of theorem (3.2.3), $h_n^{(1)}(x)$ has a limiting normal distribution with mean h(x) and variance $\frac{\alpha_n}{n} \frac{h(x)}{1-F(x)}$. Therefore, $h_n^{(1)}(x)$ is asymptotically unbiased estimator with asymptotic variance $\frac{\alpha_n}{n} \frac{h(x)}{1-F(x)}$. By the asymptotic equivalence in section 3.1, all three estimators have the same limiting normal distribution. As such the asymptotic unbiasedness of $h_n^{(2)}(x)$ and $h_n^{(3)}(x)$ also follows from this important theorem (3.2.3). However, one should beware that the choose of $\delta_n(x) = w_n(x)$ and the requirement that the weight function w has finite support are required for the asymptotic equivalence.

Remark: Note that all three estimators have the same asymptotic variance, $\frac{\alpha_n}{n}\frac{h(x)}{1-F(x)}$. If we particularly assume $\alpha_n = o(n)$

$$\frac{\alpha_n}{n}\frac{h(x)}{1-F(x)} \longrightarrow 0 \text{ as } n \longrightarrow \infty$$

(i.e. all three estimators are consistent). In fact, the variance converges in exactly the same way as α_n/n.

This chapter has introduced quite a few conditions. Therefore, examples of $\delta_n(x)$ should be constructed to show how these conditions are satisfied.

Example 3.2.1:

Suppose that $\delta_n(x) = \sqrt{n}\ w(\sqrt{n}\ x)$ where

$$w(x) = \begin{cases} \frac{1}{2}, & |x| \le 1 \\ 0, & |x| > 1 \end{cases}. \tag{3.2.2}$$

Here $\delta_n(x)$ is a particular example of (1.1.2) with $b_n = 1/\sqrt{n}$. Clearly, w has finite support. Hence, if the life distribution F is continuously differentiable with density f, all three estimators are asymptotic equivalent by theorem (3.1.1).

We now turn our attention to the conditions inside the asymptotic normality theorem (3.2.2). Since

$$\alpha_n = \int_{-\infty}^{-\infty} \delta_n^2(x)dx$$

$$= \int_{-\infty}^{-\infty} [\sqrt{n}\ w(\sqrt{n}\ x)]^2 dx$$

$$= \int_{-1/\sqrt{n}}^{1/\sqrt{n}} \left[\frac{\sqrt{n}}{2} \right]^2 dx$$

$$= \frac{n}{4} \times \frac{2}{\sqrt{n}}$$

$$= \frac{\sqrt{n}}{2}.$$

Hence, $\alpha_n = o(n)$. Now let

$$\gamma_n = \int_{-\infty}^{-\infty} |\delta_n(x)|^{2+\eta}\ dx, \qquad \eta > 0$$

$$= \int_{-1/\sqrt{n}}^{1/\sqrt{n}} \left[\frac{\sqrt{n}}{2} \right]^{2+\eta} dx$$

$$= \left[\frac{\sqrt{n}}{2} \right]^{2+\eta} \left[\frac{2}{\sqrt{n}} \right]$$

$$= \left[\frac{\sqrt{n}}{2} \right]^{1+\eta} .$$

Thus,

$$\frac{\gamma_n}{n^{\eta/2} \alpha_n^{1+\eta/2}} = \frac{(\sqrt{n}/2)^{1+\eta}}{n^{\eta/2} (\sqrt{n}/2)^{1+\eta/2}}$$

$$= \frac{1}{(2\sqrt{n})^{\eta/2}} \longrightarrow 0 \text{ as } n \longrightarrow \infty.$$

We then assume f is continuously differentiable. So

$$E[f_n(x)] = \int_{-\infty}^{-\infty} \delta_n(x-u)\ f(u)du$$

$$= \int_{-\infty}^{-\infty} \sqrt{n}\ w(\sqrt{n}(x-u))f(u)du$$

$$= \int_{-\infty}^{-\infty} w(v)\ f(x-\sqrt{n}\ v)dv$$

$$= f(x) - \frac{1}{\sqrt{n}} f'(u) \int_{-\infty}^{-\infty} w(v)vdv + o\left[\frac{1}{\sqrt{n}} \right]$$

$$= f(x) + o\left[\frac{1}{\sqrt{n}} \right].$$

Therefore,

$$\left[\frac{n}{\alpha_n} \right]^{\frac{1}{2}} \{E[f_n(x)] - f(x)\}$$

$$= \left[\frac{2n}{\sqrt{n}} \right]^{\frac{1}{2}} \left\{ f(x) + o\left(\frac{1}{\sqrt{n}}\right) - f(x) \right\}$$

$$= 2\sqrt{n} \left[o\left(\frac{1}{\sqrt{n}} \right) \right] \longrightarrow 0 \text{ as } n \longrightarrow \infty.$$

Hence, by assuming f is continuously differentiable and using (3.2.2), we have the asymptotic normality for all three estimators. Furthermore, as $\alpha_n = o(n)$, all three estimators are consistent.

3.3 GLOBAL MEASURE OF DEVIATION

Rice and Rosenblatt (1976) strengthened the results of Bickel and Rosenblatt (1973), for the sample density function to obtain the asymptotic global results for $h_n^{(1)}$ by a weak approximation. Here I will discuss their result in detail.

Unless otherwise specified, we assume δ_n as example (1.1.2) throughout this entire section that is,

$$\delta_n = \frac{1}{b_n} w\left[\frac{x}{b_n} \right].$$

Let $M_n = \max_{|t| \le \alpha(n)} |[nb_n f^{-1}(t)]^{\frac{1}{2}} (f_n(t)-f(t))|$, where $\alpha(n) \to \infty$ as $n \to \infty$ but $\log \alpha(n) = O(n)$. The weight function w is assumed to be zero outside an interval [–A,A] and either (a) absolutely continuous on [–A,A] or (b) absolutely continuous on $(-\infty,\infty)$ with derivative w' such that $\int_{-\infty}^{\infty} |w'(t)|^k < \infty$, $k = 1,2$. Then the following theorem can be shown to hold.

<u>Theorem 3.3.1</u> (Rice and Rosenblatt 1976):

Let f be a positive density on $(-\infty,\infty)$ that is twice continuously differentiable with a bounded second derivative.

Set

$$b_n = n^{-\delta}, \quad o < \delta < \frac{1}{2} .$$

Choose the sequence $\alpha(n)$ so that

$$\sup_{|t| \le \alpha(n)} f^{-\frac{1}{2}}(t), \sup_{|t| \le \alpha(n)} \frac{f'(t)}{f(t)} = O(n^{+\frac{1}{2}(1-\delta-\epsilon)}), O(n^{+\frac{1}{2}(\delta-\epsilon)})$$

for some $\epsilon > o$ with $1 - \delta - \epsilon$, $\delta - \epsilon > o$ as $n \to \infty$. Let $c(n) = 2\alpha(n)/b_n$.

Then $$P\left\{(2 \log c(n))^{\frac{1}{2}} \left[\frac{M_n}{(\lambda(w))^{\frac{1}{2}}} - d_n \right] < x \right\} \to e^{-2e^{-x}},$$

where $$\lambda(w) = \int_{-\infty}^{\infty} w^2(t)dt,$$

and

$$d_n = (2 \log c(n))^{\frac{1}{2}} + \frac{1}{(2 \log c(n))^{\frac{1}{2}}} \left\{ \log \frac{K_1(w)}{\pi^{\frac{1}{2}}} + \frac{1}{2} \log \log c(n) \right\},$$

with $\quad K_1(w) = \frac{w^2(A) + w^2(-A)}{2\lambda(w)}$

if $K_1(w) > 0$ and otherwise

$$d_n = (2 \log c(n))^{\frac{1}{2}} + \frac{1}{(2 \log c(n))^{\frac{1}{2}}} \left[\log \frac{1}{\pi} \frac{K_2(w)}{2} \right]$$

where $\quad K_2(w) = \left\{ \frac{\int_{-\infty}^{\infty} [w'(t)]^2 dt}{\lambda(w)} \right\}^{\frac{1}{2}}.$

<u>Proof:</u>

$$E[f_n(t)] = \int_{-\infty}^{\infty} \frac{1}{b_n} w\left[\frac{x-u}{b_n} \right] f(u)du$$

$$= \int_{-\infty}^{\infty} w(v)\ f(x - b_n v)du$$

$$= \int_{-\infty}^{\infty} w(v)\ [f(x) - b_n f'(x)v + b_n{}^2 f''(x)v^2 + o(b_n^2)]dv$$

(By Taylor's Formula)

$$= f(x) + b_n^2 f''(x) \int_{-\infty}^{\infty} w(v)v^2 dv + o(b_n^2).$$

By assumption, f has a bounded second derivative and w has a finite support. Therefore,

$$\sup_t |f(t) - E[f_n(t)]| = O(b_n^2).$$

Then $[nb_n f^{-1}(t)]^{\frac{1}{2}} (f_n(t) - f(t))$ can be replaced by

$$Y_n(t) = [nb_n f^{-1}(t)]^{\frac{1}{2}} (f_n(t) - E[f_n(t)])$$

$$= b_n^{\frac{1}{2}} f^{-\frac{1}{2}}(t) \int_{-\infty}^{\infty} \frac{1}{b_n} w \left[\frac{t-s}{b_n} \right] \sqrt{n}\, d(F_n(s) - F(s)).$$

Let $Z_n^0(t) = n^{\frac{1}{2}}(F_n^*(t) - t)$ and $F_n^* = F_n(F^{-1})$ is the empirical distribution function of $F(X_1),\ldots,F(X_n)$, then

$$Y_n(t) = b_n^{-\frac{1}{2}} f^{-\frac{1}{2}}(t) \int_{-\infty}^{\infty} w \left[\frac{t-s}{b_n} \right] dZ_n^0(F(s)).$$

Let $Z^0(\cdot)$ be the Brownian bridge, that is,

$$Z^0(t) = Z(t) - tZ(1),$$

where Z is a standard Wiener process on [0,1]. The process ${}_0Y_n$, ${}_1Y_n$, ${}_2Y_n$ and ${}_3Y_n$ are given by

$${}_0Y_n = b_n^{-\frac{1}{2}} f^{-\frac{1}{2}}(t) \int_{-\infty}^{\infty} w \left[\frac{t-s}{b_n} \right] dZ^0(F(s)),$$

$$ {}_1Y_n = b_n^{-\frac{1}{2}} f^{-\frac{1}{2}}(t) \int_{-\infty}^{\infty} w\left[\frac{t-s}{b_n}\right] dZ(F(s)), $$

$$ {}_2Y_n = [b_n f(t)]^{-\frac{1}{2}} \int_{-\infty}^{\infty} w\left[\frac{t-s}{b_n}\right] (f(s))^{\frac{1}{2}} dZ(s), $$

$$ {}_3Y_n = b_n^{-\frac{1}{2}} \int_{-\infty}^{\infty} w\left[\frac{t-s}{b_n}\right] dZ(s). $$

By a theorem of Komlós, Major and Tusnády (1975) versions of Z_n^0 and Z^0 can be constructed on the same probability space so that the maximal difference

$$ \|Z_n^0 - Z^0\| = O_p(n^{-\frac{1}{2}} \log n). $$

Using integration by parts with $u = w\left[\frac{t-s}{b_n}\right]$ and $v = dZ_n^0(F(s))$

$$ Y_n(t) = [b_n f(t)]^{-\frac{1}{2}} \left\{ w\left[\frac{t-s}{b_n}\right] Z_n^0(F(s))\Big|_{t-Ab_n}^{t+Ab_n} + \frac{1}{b_n} \int_{-\infty}^{\infty} Z_n^0(F(s)) w'\left[\frac{t-s}{b_n}\right] ds \right\} $$

$$ = [b_n f(t)]^{-\frac{1}{2}} \left\{ w(-A) Z_n^0(F(+Ab_n)) - w(A) Z_n^0(F(t-Ab_n)) + \frac{1}{b_n} \int_{-\infty}^{\infty} Z_n^0(F(s)) w'\left[\frac{t-s}{b_n}\right] ds \right\}. $$

The first two terms inside the curly brackets are 0 in the event of assumption (b) holds but (a) does not. Hence,

$$Y_n(t) = b_n^{-\frac{3}{2}} f^{-\frac{1}{2}}(t) \int_{-\infty}^{\infty} Z_n^0(F(s))w'\left[\frac{t-s}{b_n}\right] ds$$

$$= b_n^{-\frac{1}{2}} f^{-\frac{1}{2}}(t) \int_{-\infty}^{\infty} Z_n^0(F(t-b_n u))w'(u)du,$$

and

$${}_0Y_n(t) = b_n^{-\frac{1}{2}} f^{-\frac{1}{2}}(t) \int_{-\infty}^{\infty} Z^0(F(t-b_n u))w'(u)du.$$

Therefore,

$$|Y_n(t) - {}_0Y_n(t)| = |b_n^{-\frac{1}{2}} f^{-\frac{1}{2}}(t)| \left|\int_{-\infty}^{\infty} [Z_n^0(F(t-b_n u)) - Z^0(F(t-b_n u))]w'(u)du\right|$$

$$\leq |f^{-\frac{1}{2}}(t)| b_n^{-\frac{1}{2}} \sup_u |Z_n^0(F(u)) - Z^0(F(u))| \int_{-\infty}^{\infty} |w'(u)|du$$

$$= |f^{-\frac{1}{2}}(t)| b_n^{-\frac{1}{2}}\, O_p(n^{-\frac{1}{2}} \log n)$$

$$= |f^{-\frac{1}{2}}(t)|\, O_p(b_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log n).$$

Hence $\sup_{|t| \leq \alpha(n)} |Y_n(t) - {}_0Y_n(t)| = O_p(b_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log n) \sup_{|t| \leq \alpha(n)} f^{-\frac{1}{2}}(t)$

(3.3.1)

Since $Z^0(t) = Z(t) - tZ(1)$, then

$$|{}_0Y_n(t) - {}_1Y_n(t)|$$

$$= \left| b_n^{-\frac{1}{2}} f^{-\frac{1}{2}}(t) \int_{-\infty}^{\infty} w\left[\frac{t-s}{b_n} \right] d(Z^0(F(s)) - Z(F(s))) \right|$$

$$= b_n^{-\frac{1}{2}} f^{-\frac{1}{2}}(t) \left| \int_{-\infty}^{\infty} w\left[\frac{t-s}{b_n} \right] d(Z(F(s)) - F(s)Z(1) - Z(F(s))) \right|$$

$$= b_n^{-\frac{1}{2}} f^{-\frac{1}{2}}(t) |Z(1)| \left| \int_{-\infty}^{\infty} w\left[\frac{t-s}{b_n} \right] dF(s) \right|$$

$$= b_n^{\frac{1}{2}} f^{-\frac{1}{2}}(t) |Z(1)| \; |E(f_n(t))|$$

$$= O_p(1) \; b_n^{\frac{1}{2}} f^{-\frac{1}{2}}(t).$$

Therefore,

$$\sup_{|t| \leq \alpha(n)} |{}_0Y_n(t) - {}_1Y_n(t)| = O_p(1) \; b_n^{\frac{1}{2}} \sup_{|t| \leq \alpha(n)} f^{-\frac{1}{2}}(t) \tag{3.3.2}$$

Now we are going to show that the process ${}_1Y_n$ and ${}_2Y_n$ have the same probability structure.

$$E[{}_1Y_n(t)] = b_n^{-\frac{1}{2}} f^{-\frac{1}{2}}(t) \int_{-\infty}^{\infty} w\left[\frac{t-s}{b_n} \right] dE[Z(F(s))]$$

$$= 0$$

and

$$E[{}_2Y_n(t)] = b_n^{-\frac{1}{2}} f^{-\frac{1}{2}}(t) \int_{-\infty}^{\infty} w\left[\frac{t-s}{b_n} \right] (f(s))^{\frac{1}{2}} \, dE[Z(s)]$$

$$= 0.$$

For $0 < t_1, t_2 < \infty$,

$E[{}_1Y_n(t_1) \, {}_1Y_n(t_2)]$

$$= E\left\{\left[b_n^{-\frac{1}{2}} f^{-\frac{1}{2}}(t_1) \int_{-\infty}^{\infty} w\left[\frac{t_1-s_1}{b_n}\right] dZ(F(s_1))\right]\left[b_n^{-\frac{1}{2}} f^{-\frac{1}{2}}(t_2) \int_{-\infty}^{\infty} w\left[\frac{t_2-s_2}{b_n}\right] dZ(F(s_2))\right]\right\}$$

$$= b_n^{-1} f^{-\frac{1}{2}}(t_1) f^{-\frac{1}{2}}(t_2) \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} w\left[\frac{t_1-s_1}{b_n} \right] w\left[\frac{t_2-s_2}{b_n} \right] \, d \, E[Z(F(s_1))Z(F(s_2))]$$

$$= b_n^{-1} f^{-\frac{1}{2}}(t_1) f^{-\frac{1}{2}}(t_2) \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} w\left[\frac{t_1-s_1}{b_n} \right] w\left[\frac{t_2-s_2}{b_n} \right] \, d(F(s_1) \wedge F(s_2))$$

$$= b_n^{-1} f^{-\frac{1}{2}}(t_1) f^{-\frac{1}{2}}(t_2) \int_{-\infty}^{\infty} w\left[\frac{t_1-s}{b_n} \right] w\left[\frac{t_2-s}{b_n} \right] dF(s)$$

and

$E[{}_2Y_n(t_1) {}_2Y_n(t_2)]$

$$= E\left\{\left[b_n^{-\frac{1}{2}} f^{-\frac{1}{2}}(t_1) \int_{-\infty}^{\infty} w\left[\frac{t_1-s_1}{b_n} \right] (f(s_1))^{\frac{1}{2}} \, dZ(s_1) \right].\right.$$

$$\left[b_n^{-\frac{1}{2}} f^{-\frac{1}{2}}(t_2)\int_{-\infty}^{\infty} w\left[\frac{t_2-s_2}{b_n}\right] (f(s_2))^{\frac{1}{2}}\, dZ(s_2)\right]\Bigg\}$$

$$= b_n^{-1} f^{-\frac{1}{2}}(t_1) f^{-\frac{1}{2}}(t_2)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} w\left[\frac{t_1-s_1}{b_n}\right] w\left[\frac{t_2-s_2}{b_n}\right](f(s_1))^{\frac{1}{2}}(f(s_2))^{\frac{1}{2}}\, d[E(Z(s_1)Z(s_2)]$$

$$= b_n^{-1} f^{-\frac{1}{2}}(t_1) f^{-\frac{1}{2}}(t_2)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} w\left[\frac{t_1-s_1}{b_n}\right] w\left[\frac{t_2-s_2}{b_n}\right](f(s_1))^{\frac{1}{2}}(f(s_2))^{\frac{1}{2}}\, d(s_1 \wedge s_2)$$

$$= b_n^{-1} f^{-\frac{1}{2}}(t_1) f^{-\frac{1}{2}}(t_2)\int_{-\infty}^{\infty} w\left[\frac{t_1-s}{b_n}\right] w\left[\frac{t_2-s}{b_n}\right] f(s)ds$$

Hence, ${}_1Y_n$ and ${}_2Y_n$ have the same mean and covariance implying that they have same probability structure. Further,

$${}_2Y_n(t) - {}_3Y_n(t) = b_n^{-\frac{1}{2}} \int_{-\infty}^{\infty} w\left[\frac{t-s}{b_n}\right] \left(\left[\frac{f(s)}{f(t)}\right]^{\frac{1}{2}} -1\right) dZ(s).$$

Using integration by parts with $u = w\left[\frac{t-s}{b_n}\right] \left(\left[\frac{f(s)}{f(t)}\right]^{\frac{1}{2}} -1\right)$ and $v = dZ(s)$, then

$$|{}_2Y_n(t) - {}_3Y_n(t)|$$

$$= \left| b_n^{-\frac{1}{2}} w\left[\frac{t-s}{b_n}\right] \left(\left[\frac{f(s)}{f(t)}\right]^{\frac{1}{2}} -1\right) Z(s) \Big|_{t-Ab_n}^{t+Ab_n}\right.$$

$$+ \; b_n^{-\frac{3}{2}} \int_{-\infty}^{\infty} Z(s) w'\left[\frac{t-s}{b_n} \right] \left[\left[\frac{f(s)}{f(t)} \right]^{\frac{1}{2}} - 1 \right] ds$$

$$+ \; \frac{b_n^{-\frac{1}{2}}}{2} \int_{-\infty}^{\infty} Z(s) w\left[\frac{t-s}{b_n} \right] \left[\frac{f'(s)}{[f(s)f(t)]^{\frac{1}{2}}} \right] ds \; \Bigg|$$

$$= \Bigg| \; b_n^{-\frac{1}{2}} \, w(-A) \left[\left[\frac{f(t+Ab_n)}{f(t)} \right]^{\frac{1}{2}} - 1 \right] Z(t+Ab_n)$$

$$- \; b_n^{-\frac{1}{2}} \, w(A) \left[\left[\frac{f(t+Ab_n)}{f(t)} \right]^{\frac{1}{2}} - 1 \right] Z(t-Ab_n)$$

$$+ \; b_n^{-\frac{3}{2}} \sup_{s} Z(s) \sup_{|s-t|<Ab_n} \left[\left[\frac{f(s)}{f(t)} \right]^{\frac{1}{2}} - 1 \right] \int_{t-Ab_n}^{t+Ab_n} w'\left[\frac{t-s}{b_n} \right] ds$$

$$+ \; b_n^{-\frac{1}{2}} \sup_{s} Z(s) \sup_{|s-t|<Ab_n} \frac{f'(s)}{[f(s)f(t)]^{\frac{1}{2}}} \int_{t-Ab_n}^{t+Ab_n} w\left[\frac{t-s}{b_n} \right] ds \; \Bigg| .$$

The first two terms inside the absolute value sign vanish in the event that assumption (b) holds but (a) does not. Hence

$$|\,{}_2Y_n(t) - {}_3Y_n(t)\,|$$

$$= b_n^{-\frac{1}{2}} \, O_p(1) \; O(b_n) + b_n^{\frac{1}{2}} \, O_p(1) \sup_{|s-t|<Ab_n} \frac{f'(s)}{[f(s)f(t)]^{\frac{1}{2}}}$$

$$= O_p(b_n^{\frac{1}{2}}) \sup_{|s-t|<Ab_n} \frac{f'(s)}{[f(s)f(t)]^{\frac{1}{2}}} .$$

When $n \longrightarrow \infty$, $b_n \downarrow 0$ and

$$\sup_{|s-t|<Ab_n} \frac{f'(s)}{[f(s)f(t)]^{\frac{1}{2}}} \longrightarrow \frac{f'(t)}{f(t)} .$$

By assumption,

$$\sup_{|t|\leq\alpha(n)} \frac{f'(t)}{f(t)} = O(n^{\frac{1}{2}(\delta-\epsilon)}).$$

Therefore,

$$\begin{aligned} \sup_{|t|\leq\alpha(n)} |{}_2Y_n(t)-{}_3Y_n(t)| &= O_p(b_n^{\frac{1}{2}})\ O(n^{\frac{1}{2}(\delta-\epsilon)}) \\ &= O_p(n^{-\frac{\delta}{2}})\ O(n^{\frac{1}{2}(\delta-\epsilon)}), \qquad \text{by assumption} \\ &= O(n^{-\frac{1}{2}\epsilon}). \qquad (3.3.3) \end{aligned}$$

Finally, we try to indicate that ${}_3Y_n$ has the same probability structure as $Y(t)$ where $Y(t)$ is the Gaussian process

$$Y(t) = \int_{-\infty}^{\infty} w(t-s)\ dZ(s).$$

$$E[{}_3Y_n(t)] = b_n^{-\frac{1}{2}} \int_{-\infty}^{\infty} w\left[\frac{t-s}{b_n}\right] dE[Z(s)]$$

$$= 0$$

and

$$E[Y(t)] = \int_{-\infty}^{\infty} w(t-s)dE[Z(s)]$$

$$= 0.$$

For $0 < t_1, t_2 < \infty$,

$$E[{}_3Y_n(t_1)\,{}_3Y_n(t_2)]$$

$$= E\left\{\left[b_n^{-\frac{1}{2}}\int_{-\infty}^{\infty} w\left[\frac{t_1-s_1}{b_n}\right] d\,Z(s_1)\right]\left[b_n^{-\frac{1}{2}}\int_{-\infty}^{\infty} w\left[\frac{t_2-s_2}{b_n}\right] dZ(s_2)\right]\right\}$$

$$= b_n^{-1}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} w\left[\frac{t_1-s_1}{b_n}\right] w\left[\frac{t_2-s_2}{b_n}\right] d\ E[Z(s_1)Z(s_2)]$$

$$= b_n^{-1}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} w\left[\frac{t_1-s_1}{b_n}\right] w\left[\frac{t_2-s_2}{b_n}\right] d(s_1 \wedge s_2)$$

$$= b_n^{-1}\int_{-\infty}^{\infty} w\left[\frac{t_1-s}{b_n}\right] w\left[\frac{t_2-s}{b_n}\right] ds$$

and

$$E[Y(t_1)Y(t_2)]$$

$$= E\left\{\left[\int_{-\infty}^{\infty} w(t_1-s_1)d\ Z(s_1)\right]\left[\int_{-\infty}^{\infty} w(t_2-s_2)dZ(s_2)\right]\right\}.$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} w(t_1-s_1)w(t_2-s_2)dE[Z(s_1)Z(s_2)]$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} w(t_1-s_1)w(t_2-s_2)d(s_1 \wedge s_2)$$

$$= \int_{-\infty}^{\infty} w(t_1-s)w(t_2-s)ds$$

$$= b_n^{-1}\int_{-\infty}^{\infty} w\left[\frac{t_1-s}{b_n}\right] w\left[\frac{t_2-s}{b_n}\right] ds.$$

Since ${}_3Y_n$ and Y have the same mean and covariance implying that they have same probability structure. By all the above estimates, we have indeed shown that the limiting distribution of M_n is the same as that of $\sup_{|t|\leq\alpha(n)} Y(t)$. Applying the known result on the maxima of stationary Gaussian process in Bickel and Rosenblatt (1973) leads to the conclusion of this theorem.

Now let

$$\overline{M}_n = \max_{|t|\leq\alpha(n)} |(nb_n f^{-1}(t))^{\frac{1}{2}} (1-F(t)) (h_n^{(1)}(t)-h(t))|.$$

The global result of $h_n^{(1)}$ is a direct corollary of theorem (3.3.1).

<u>Corollary 3.3.1</u> (Rice and Rosenblatt 1976):

Under the condition of theorem (3.3.1) and the additional assumption

$$\sup_{|t| \le \alpha(n)} (1-F(t))^{-1} = o(n^{\frac{1}{2}}) \tag{3.3.4}$$

one has

$$P\left\{(2 \log c(n))^{\frac{1}{2}}\left[\frac{M_n}{(\lambda(w))^{\frac{1}{2}}} - d_n\right] < x\right\} \longrightarrow e^{-2e^{-x}}$$

as $n \longrightarrow \infty$.

<u>Proof</u>:

What we need to prove here is $\overline{M}_n \longrightarrow M_n$ as $n \longrightarrow \infty$.

$$\overline{M}_n = \max_{|t| \le \alpha(n)} |(nb_nf^{-1}(t))^{\frac{1}{2}}(1-F(t))(h_n^{(1)}(t)-h(t))|$$

$$= \max_{|t| \le \alpha(n)} \Bigg|(nb_nf^{-1}(t))^{\frac{1}{2}}(1 - F(t)) \left\{ \frac{f_n(t)-f(t)}{1-F(t)} + f_n(t)\left[\frac{1}{1-F_n(t)} - \frac{1}{1-F(t)}\right]\right\}\Bigg|$$

$$= \max_{|t| \le \alpha(n)} |(nb_nf^{-1}(t))^{\frac{1}{2}}| \max_{|t| \le \alpha(n)} \Bigg|f_n(t)-f(t) + f_n(t)\left\{\sum_{j=0}^{\infty}(-1)^j\left[\frac{F(t)-F_n(t)}{1-F(t)}\right]^j -1\right\}\Bigg|$$

$$= \max_{|t| \le \alpha(n)} |(nb_nf^{-1}(t))^{\frac{1}{2}}| \max_{|t| \le \alpha(n)} \Bigg|f_n(t)-f(t) \cdot + f_n(t)\left[O_p(n^{-\frac{1}{2}})\frac{1}{1-F(t)}\right]\Bigg|$$

$$= \max_{|t| \le \alpha(n)} |(nb_n f^{-1}(t))^{\frac{1}{2}}| \max_{|t| \le \alpha(n)} \Big| f_n(t) - f(t) + f_n(t) \left[O_p(n^{-\frac{1}{2}}) \; o(n^{\frac{1}{2}}) \right] \Big|$$

$$= \max_{|t| \le \alpha(n)} |(nb_n f^{-1}(t))^{\frac{1}{2}}| \max_{|t| \le \alpha(n)} \Big| f_n(t) - f(t) + f_n(t) \; o(1) \Big|$$

$$= \max_{|t| \le \alpha(n)} |(n \; b_n f^{-1}(t))^{\frac{1}{2}} \; (f_n(t) - f(t)) \; | + o(1)$$

$$= M_n + o(1).$$

Hence, the corollary.

Sethuraman and Singpurwalla (1981) have further obtained, in much the same way as Bickel and Rosenblatt, the asymptotic global result for $h_n^{(3)}(x)$:

<u>Theorem 3.3.2</u>:

Assume the following conditions hold:

(A1) w has bandwidth $2b_n A$,

(A2) $B''(x)$ is bounded on $0 \le x \le K$ and $\inf_{0 \le x \le K+A} B'(x) > 0$

where $B(x) = \frac{F(x)}{1-F(x)}$ and $K + A < X^{(n)}$, (A3) let $\lambda(w) = \int_{-\infty}^{\infty} w^2(t)dt$, then either (a) $\int_{-\infty}^{\infty} |w'(t)|dt < \infty$ and

$K_1(w) = \frac{w^2(A)+w^2(-A)}{2\lambda(w)} > 0$ or (b) $\int_{-\infty}^{\infty} (w'(t))^2 dt < \infty$ and

$K_1(w) = 0$,

(A4) $h(x)$ is twice continuously differentiable,

(A5) $nb_n^5 \log b_n \to 0$ as $n \to \infty$.

Let

$$M_n = \max_{b_n A \le x \le L} \left| \frac{\sqrt{nb_n}}{\sqrt{B'(x)}} (h_n^{(3)}(x) - h(x)) \right| ,$$

$$C_n = (2 \log (K/b_n))^{\frac{1}{2}},$$

$$\beta_n = C_n/(\lambda(w))^{\frac{1}{2}},$$

and

$$\alpha_n = \begin{cases} (\lambda(w))^{\frac{1}{2}}[C_n + \log (C_n K_1(w)/\sqrt{2\pi})]/C_n & , \text{ when } K_1(w) > 0 \\ (\lambda(w))^{\frac{1}{2}}[C_n + \log (K_2(w)/\pi)]/C_n, & \text{when } K_1(w) = 0. \end{cases}$$

Then for $0 < x < \infty$,

$$P[\beta_n(M_n - \alpha_n) \le z] \to e^{-2e^{-z}}.$$

The proof of theorem (3.3.2) is given in Sethuraman and Singpurwalla (1981) and will not be presented here.

Towards this end, one should be reminded of the previous important result, asymptotic equivalence, in section 3.1. That is the two aforestated theorems in this section should both hold for all three estimators by their asymptotic equivalence.

CHAPTER IV
SIMULATIONS AND APPLICATIONS

In this Chapter, I report the results of some numerical experiments conducted to demonstrate how reasonable the following results are for a finite sample size:

(1) the estimators are asymptotic equivalent,

(2) the asymptotic variance formula for the estimator is adequate,

(3) the estimators are asymptotically normal.

Furthermore, the performances of these estimators were compared by using the average square error.

4.1 PROCEDURE

To study the asymptotic properties of the estimators mentioned above, it seemed best to draw samples from populations with smooth hazard functions. This was reminiscent of the well known exponential life distribution. We therefore used the GGEXN package of IMSL (International Mathematical and Statistical Library) to generate 1000 samples of size 100 with X exponentially distributed, mean one. To test the asymptotic results, the same was done with samples of size 500 and 1000. As unit mean was used, we had unit hazard rate here disregarding what the value of x was.

As regard to the δ_n–function, example (2.2.1) was used here. That is, $\delta_n = \sqrt{n}w(\sqrt{n}x)$ where

$$w(x) = \begin{cases} \frac{1}{2} & , |x| \le 1 \\ 0 & , |x| > 1 \end{cases} .$$

I used (3.2.2) because it satisfies all the conditions of the previous asymptotic theorems. Unless otherwise specified, n is the sample size in this Chapter.

To measure the goodness of performance of an estimator, many authors use the M.I.S.E. (Mean Integrated Square Error). However, the M.I.S.E. is quite difficult to obtain numerically. I therefore decided to use the average square error being defined as follows:

$$\left[\frac{1}{1000} \right] \sum_{i=1}^{1000} [h_n^{(j)}(x) - h(x)]^2, \qquad j = 1,2,3.$$

From Chapter 3, I noted that all three estimators have asymptotic variance given by

$$\mathrm{Var}\ (h_n^{(i)}(x)) \sim \frac{\alpha_n}{n} \frac{h(x)}{1-F(x)}\ ,\ i = 1,2,3.$$

Hence, the asymptotic variance here is

$$\mathrm{Var}\ (h_n^{(i)}(x)) \sim \frac{\sqrt{n}}{2n} \frac{1}{e^{-x}} = \frac{1}{2\sqrt{n}}\, e^{x} \qquad (4.2.1)$$

To test the adequacy of this formula, we calculate the sample variance to compare with.

For each sample size and each estimator, the values of average square error, sample variance s^2 and estimated variance $\hat{\sigma}^2$ (4.2.1) were computed and are found in Tables 4.2.1–9.

To check the asymptotic equivalence, I calculated the average square difference between estimators,

$$\left[\frac{1}{1000}\right] \sum_{i=1}^{1000} [h_n^{(j)}(x) - h_n^{(k)}(x)]^2, \; j \neq k$$

at the sample points. Such values were computed and are given in Tables 4.2.10–12.

To compare all the above calculated values as a whole, I then computed for each sample size (i) the total average square error, (ii) the total average square difference between estimators and (iii) the relative total square difference between s^2 and $\hat{\sigma}^2$ as follows:

$$\text{(i)} \qquad \sum_{x=0.1}^{2.0} \sum_{i=1}^{1000} [h_n^{(j)}(x) - h(x)]^2, \qquad j = 1,2,3$$

$$\text{(ii)} \qquad \sum_{x=0.1}^{2.0} \sum_{i=1}^{1000} [h_n^{(j)}(x) - h(x)]^2, \qquad j \neq k$$

$$\text{(iii)} \qquad \sum_{x=0.1}^{2.0} [s^2 - \hat{\sigma}^2]^2 \, \frac{(1/2\sqrt{100})e^x}{(1/2\sqrt{n})e^x}$$

$$= \sum_{x=0.1}^{2.0} \sqrt{\frac{n}{100}} \, [s^2 - \hat{\sigma}^2]^2 .$$

Remark: The relative value of total square difference was used here because the estimated variance is a decreasing function and I wanted to compare the results of large sample sizes with samples of 100.

All these solutions are given in Tables 4.2.13–15. These computations were done with computer algorithms written in APL (A Programming Language) because of programming simplicity.

To assess asymptotic normality, I used the well known Kolmogorov–Smirnov Test for which the cumulative distribution functions for the observed data and the theoretical distribution are computed and subtracted. The Kolmogorov– Smirnov Z is determined from the largest difference (positive or negative). The larger the value of Z, the less likely it is that the observed and theoretical distributions are the same. I performed these tests with SPSS (Statistical Package for Social Sciences) using unit mean and $\hat{\sigma}^2$ as the required input information. The Kolmogorov–Smirnov Z values and 2–tailed p–values for each case are found in Tables 4.2.16–18.

4.2 CONCLUSIONS

Given that the total average square error decreases with increasing sample size (Table 4.2.13), all estimators will perform better with larger samples. $H_n^{(3)}$ gave the least total average square errors for each sample size: I therefore consider $h_n^{(3)}$ as a better estimator for finite sample sizes.

Tables 4.2.14–16 indicate that the total average square difference between the estimators decreases as sample size increases. When $n = 1000$, this difference is almost negligible. This result confirms the asymptotic equivalence of the estimators. The tables also show that the total average square difference between $h_n^{(2)}$ and $h_n^{(3)}$ is much smaller than the other pairs; leading me to conclude that $h_n^{(2)}$ and $h_n^{(3)}$ tend to each other faster than any other pair of estimators.

Tables 4.2.1–9 indicate that the average square error and sample variance are increasing functions of x, as would be expected given that the estimated variance $\hat{\sigma}^2((1/2\sqrt{n})e^x)$ is also an increasing function of x. Moreover, the relative total square difference between the s^2 and $\hat{\sigma}^2$ decreases as n increases (Tables 4.2.15), demonstrating that the asymptotic variance formula is quite adequate.

I finally turn to asymptotic normality. Tables 4.2.16–18 show that when sample size increases, the normality hypothesis is rejected at a Type I error rate of 5% level of significance less frequently. When n = 1000, all the hypotheses at different values of x are accepted for $h_n^{(1)}$ and only 1 and 2 are being rejected for $h_n^{(2)}$ and $h_n^{(3)}$ respectively. Hence, if we further increase the sample size to 5000 or 10,000, a better result should be obtained; however, such a test could not be performed due to computer storage problems. Nevertheless, the Kolmogorov–Smirnov Tests indicate the asymptotic normality of the three estimators.

TABLE 4.2.1

THE TABLE BELOW GIVES THE RESULTS FOR ESTIMATOR

$H_n^{(1)}$ WHEN 1000 SAMPLES OF 100 WERE USED

x	average value	average square error	sample variance	estimated variance
0.1	0.99912	0.05848	0.05853	0.05526
0.2	1.00116	0.059	0.05905	0.06107
0.3	1.00131	0.06802	0.06808	0.06749
0.4	1.00935	0.07361	0.07359	0.07459
0.5	1.00252	0.08165	0.08173	0.08244
0.6	0.99926	0.09122	0.09131	0.09111
0.7	1.00175	0.10538	0.10548	0.10069
0.8	1.01436	0.11224	0.11215	0.11128
0.9	1.03264	0.13784	0.13691	0.12298
1.0	1.02836	0.14426	0.1436	0.13591
1.1	1.004	0.16602	0.16617	0.15021
1.2	1.01633	0.1842	0.18412	0.16601
1.3	1.03141	0.20398	0.20319	0.18346
1.4	1.00933	0.19759	0.1977	0.20276
1.5	1.01904	0.24172	0.2416	0.22408
1.6	1.02075	0.2901	0.28996	0.24765
1.7	1.02872	0.33998	0.3395	0.2737
1.8	1.02427	0.36339	0.36316	0.30248
1.9	0.9981	0.34675	0.3471	0.33429
2.0	1.03943	0.44834	0.44724	0.36945

TABLE 4.2.2

THE TABLE BELOW GIVES THE RESULTS FOR ESTIMATOR $H_n^{(2)}$ WHEN 1000 SAMPLES OF 100 WERE USED

x	average value	average square error	sample variance	estimated variance
0.1	0.99246	0.05699	0.05699	0.05526
0.2	0.99215	0.05731	0.0573	0.06107
0.3	0.99336	0.06603	0.06605	0.06749
0.4	0.99954	0.07128	0.07135	0.07459
0.5	0.9915	0.07904	0.07905	0.08244
0.6	0.98817	0.08859	0.08853	0.09111
0.7	0.98972	0.10207	0.10206	0.10069
0.8	1.00168	0.10786	0.10796	0.11128
0.9	1.0179	0.1328	0.13262	0.12298
1.0	1.01026	0.1358	0.13583	0.13591
1.1	0.98439	0.15716	0.15707	0.15021
1.2	0.99961	0.17495	0.17513	0.16601
1.3	1.00544	0.18555	0.1857	0.18346
1.4	0.98498	0.18453	0.18448	0.20276
1.5	0.99474	0.22855	0.22876	0.22408
1.6	0.98898	0.26494	0.26509	0.24765
1.7	0.99769	0.30826	0.30856	0.2737
1.8	0.97973	0.31079	0.31069	0.30248
1.9	0.96393	0.32328	0.3223	0.33429
2.0	0.99234	0.38371	0.38403	0.36945

TABLE 4.2.3

THE TABLE BELOW GIVES THE RESULTS FOR ESTIMATOR $H_n^{(3)}$ WHEN 1000 SAMPLES OF 100 WERE USED

x	average value	average square error	sample variance	estimated variance
0.1	0.98701	0.05642	0.0563	0.05526
0.2	0.98613	0.05667	0.05654	0.06107
0.3	0.98669	0.06518	0.06507	0.06749
0.4	0.99213	0.07018	0.07019	0.07459
0.5	0.98339	0.07782	0.07763	0.08244
0.6	0.97922	0.08709	0.08674	0.09111
0.7	0.97984	0.10012	0.09982	0.10069
0.8	0.99062	0.10529	0.10531	0.11128
0.9	1.00539	0.12877	0.12887	0.12298
1.0	0.99655	0.13151	0.13163	0.13591
1.1	0.9697	0.1526	0.15183	0.15021
1.2	0.98301	0.16858	0.16846	0.16601
1.3	0.98697	0.17803	0.17804	0.18346
1.4	0.96495	0.17701	0.17596	0.20276
1.5	0.97228	0.21741	0.21686	0.22408
1.6	0.96425	0.25053	0.2495	0.24765
1.7	0.96995	0.28873	0.28811	0.2737
1.8	0.95007	0.29157	0.28937	0.30248
1.9	0.9314	0.30132	0.29691	0.33429
2.0	0.95529	0.35187	0.35023	0.36945

TABLE 4.2.4

THE TABLE BELOW GIVES THE RESULTS FOR ESTIMATOR $H_n^{(1)}$ WHEN 1000 SAMPLES OF 500 WERE USED

x	average value	average square error	sample variance	estimated variance
0.1	1.01217	0.02635	0.02623	0.02471
0.2	1.00214	0.02786	0.02789	0.02731
0.3	0.99691	0.02944	0.02946	0.03018
0.4	1.00614	0.03409	0.03408	0.03336
0.5	0.997	0.04083	0.04087	0.03687
0.6	1.00698	0.04198	0.04198	0.04074
0.7	0.99675	0.04509	0.04512	0.04503
0.8	0.99211	0.04611	0.0461	0.04976
0.9	1.01351	0.0563	0.05618	0.055
1.0	1.00551	0.06526	0.06529	0.06078
1.1	1.0027	0.06479	0.06484	0.06718
1.2	1.01454	0.07912	0.07899	0.07424
1.3	1.00131	0.08857	0.08866	0.08205
1.4	1.01556	0.10045	0.10031	0.09068
1.5	1.01245	0.10515	0.1051	0.10021
1.6	1.00299	0.10768	0.10778	0.11075
1.7	1.01724	0.12944	0.12927	0.1224
1.8	1.00106	0.13953	0.13967	0.13527
1.9	1.02658	0.17047	0.16994	0.1495
2.0	1.0108	0.18122	0.18128	0.16522

TABLE 4.2.5

THE TABLE BELOW GIVES THE RESULTS FOR ESTIMATOR $H_n^{(2)}$ WHEN 1000 SAMPLES OF 500 WERE USED

x	average value	average square error	sample variance	estimated variance
0.1	1.01069	0.02618	0.02609	0.02471
0.2	1.00095	0.02783	0.02786	0.02731
0.3	0.99606	0.02944	0.02945	0.03018
0.4	1.00407	0.0339	0.03392	0.03336
0.5	0.99522	0.04087	0.04089	0.03687
0.6	1.00517	0.04167	0.04169	0.04074
0.7	0.99411	0.04476	0.04477	0.04503
0.8	0.98958	0.04568	0.04562	0.04976
0.9	1.01028	0.05551	0.05546	0.055
1.0	1.00247	0.06495	0.06501	0.06078
1.1	0.99894	0.06429	0.06436	0.06718
1.2	1.01072	0.07831	0.07828	0.07424
1.3	0.99681	0.08751	0.08759	0.08205
1.4	1.01102	0.09916	0.09914	0.09068
1.5	1.00784	0.10378	0.10382	0.10021
1.6	0.99764	0.10629	0.10639	0.11075
1.7	1.01144	0.12806	0.12805	0.1224
1.8	0.99416	0.13798	0.13809	0.13527
1.9	1.01929	0.16719	0.16698	0.1495
2.0	1.00302	0.17886	0.17903	0.16522

TABLE 4.2.6

THE TABLE BELOW GIVES THE RESULTS FOR ESTIMATOR $H_n^{(3)}$ WHEN 1000 SAMPLES OF 500 WERE USED

x	average value	average square error	sample variance	estimated variance
0.1	1.00957	0.02609	0.02603	0.02471
0.2	0.99973	0.02776	0.02779	0.02731
0.3	0.99471	0.02937	0.02937	0.03018
0.4	1.00257	0.03379	0.03381	0.03336
0.5	0.99357	0.04075	0.04075	0.03687
0.6	1.00334	0.04149	0.04152	0.04074
0.7	0.9921	0.0446	0.04459	0.04503
0.8	0.98738	0.04552	0.04541	0.04976
0.9	1.0078	0.05518	0.05517	0.055
1.0	0.99975	0.06457	0.06464	0.06078
1.1	0.99594	0.06391	0.06396	0.06718
1.2	1.00736	0.07771	0.07773	0.07424
1.3	0.99314	0.08687	0.08691	0.08205
1.4	1.0069	0.09824	0.09829	0.09068
1.5	1.00331	0.10275	0.10284	0.10021
1.6	0.99268	0.10523	0.10528	0.11075
1.7	1.00588	0.12648	0.12657	0.1224
1.8	0.98812	0.13634	0.13633	0.13527
1.9	1.01243	0.16462	0.16463	0.1495
2.0	0.99554	0.17603	0.17619	0.16522

TABLE 4.2.7

THE TABLE BELOW GIVES THE RESULTS FOR ESTIMATOR $H_n^{(1)}$ WHEN 1000 SAMPLES OF 1000 WERE USED

x	average value	average square error	sample variance	estimated variance
0.1	1.008	0.01683	0.01678	0.01747
0.2	1.00178	0.01983	0.01985	0.01931
0.3	0.99897	0.02097	0.02099	0.02134
0.4	1.00517	0.02373	0.02372	0.02359
0.5	1.00064	0.02587	0.0259	0.02607
0.6	0.99778	0.02835	0.02837	0.02881
0.7	1.00742	0.03234	0.03232	0.03184
0.8	0.99806	0.03564	0.03567	0.03519
0.9	1.00312	0.04033	0.04036	0.03889
1.0	0.99399	0.04301	0.04301	0.04298
1.1	1.01966	0.0444	0.04406	0.0475
1.2	0.99336	0.05661	0.05662	0.0525
1.3	1.00103	0.05863	0.05869	0.05802
1.4	1.01171	0.0671	0.06703	0.06412
1.5	1.00379	0.07281	0.07287	0.07086
1.6	1.00197	0.07678	0.07685	0.07831
1.7	1.01573	0.09124	0.09108	0.08655
1.8	1.02801	0.10952	0.10885	0.09565
1.9	1.00403	0.10333	0.10342	0.10571
2.0	1.01355	0.10968	0.1096	0.11683

TABLE 4.2.8

THE TABLE BELOW GIVES THE RESULTS FOR ESTIMATOR $H_n^{(2)}$ WHEN 1000 SAMPLES OF 1000 WERE USED .

x	average value	average square error	sample variance	estimated variance
0.1	1.00747	0.01686	0.01682	0.01747
0.2	1.00109	0.01986	0.01987	0.01931
0.3	0.9981	0.02085	0.02087	0.02134
0.4	1.00415	0.02362	0.02363	0.02359
0.5	0.99982	0.02585	0.02587	0.02607
0.6	0.9967	0.0282	0.02822	0.02881
0.7	1.00641	0.03218	0.03217	0.03184
0.8	0.99663	0.03563	0.03566	0.03519
0.9	1.00152	0.04021	0.04024	0.03889
1.0	0.99271	0.04301	0.043	0.04298
1.1	1.01788	0.04404	0.04377	0.0475
1.2	0.99132	0.05633	0.05631	0.0525
1.3	0.99917	0.05835	0.05841	0.05802
1.4	1.00882	0.06667	0.06665	0.06412
1.5	1.00139	0.07257	0.07264	0.07086
1.6	0.9996	0.07633	0.0764	0.07831
1.7	1.01251	0.09046	0.09039	0.08655
1.8	1.02469	0.10825	0.10775	0.09565
1.9	1.00075	0.10259	0.10269	0.10571
2.0	1.01017	0.10902	0.10903	0.11683

TABLE 4.2.9

THE TABLE BELOW GIVES THE RESULTS FOR ESTIMATOR $H_n^{(3)}$ WHEN 1000 SAMPLES OF 1000 WERE USED

x	average value	average square error	sample variance	estimated variance
0.1	1.00692	0.01684	0.0168	0.01747
0.2	1.00047	0.01983	0.01985	0.01931
0.3	0.99742	0.02082	0.02084	0.02134
0.4	1.0034	0.02358	0.02359	0.02359
0.5	0.99899	0.02581	0.02583	0.02607
0.6	0.99579	0.02815	0.02816	0.02881
0.7	1.00539	0.0321	0.03211	0.03184
0.8	0.99551	0.03556	0.03558	0.03519
0.9	1.00029	0.0401	0.04014	0.03889
1.0	0.99136	0.04291	0.04287	0.04298
1.1	1.01635	0.04386	0.04363	0.0475
1.2	0.98967	0.05617	0.05612	0.0525
1.3	0.99733	0.05813	0.05818	0.05802
1.4	1.00677	0.06636	0.06638	0.06412
1.5	0.99914	0.07223	0.0723	0.07086
1.6	0.99712	0.07594	0.07601	0.07831
1.7	1.00972	0.08988	0.08987	0.08655
1.8	1.02158	0.10743	0.10708	0.09565
1.9	0.99739	0.10187	0.10196	0.10571
2.0	1.00642	0.10812	0.10819	0.11683

TABLE 4.2.10

THE TABLE BELOW GIVES THE AVERAGE SQUARE OF DIFFERENCES BETWEEN THE THREE ESTIMATORS WHEN 1000 SAMPLES OF 100 WERE USED

x	$H_n^{(1)}$ & $H_n^{(2)}$	$H_n^{(1)}$ & $H_n^{(3)}$	$H_n^{(2)}$ & $H_n^{(3)}$
0.1	0.00072	0.00083	3.1856e-5
0.2	0.00086	0.00101	3.8862e-5
0.3	0.00086	0.00102	4.8186e-5
0.4	0.00104	0.00125	5.9851e-5
0.5	0.00112	0.00139	7.2686e-5
0.6	0.00124	0.00154	8.9704e-5
0.7	0.00145	0.00183	1.1119e-4
0.8	0.00174	0.00219	1.4051e-4
0.9	0.00206	0.00264	1.8578e-4
1.0	0.00287	0.00371	2.2469e-4
1.1	0.00287	0.00384	2.6585e-4
1.2	0.00347	0.0045	3.4978e-4
1.3	0.00445	0.00619	4.328e-4
1.4	0.00408	0.00581	5.1721e-4
1.5	0.00516	0.00723	6.8948e-4
1.6	0.00686	0.00993	8.8666e-4
1.7	0.00878	0.01273	1.1923e-3
1.8	0.01155	0.01761	1.3431e-3
1.9	0.00902	0.01348	1.721e-3
2.0	0.01794	0.02725	2.3942e-3

TABLE 4.2.11

THE TABLE BELOW GIVES THE AVERAGE SQUARE OF DIFFERENCES BETWEEN THE THREE ESTIMATORS WHEN 1000 SAMPLES OF 500 WERE USED

x	$H_n^{(1)}$ & $H_n^{(2)}$	$H_n^{(1)}$ & $H_n^{(2)}$	$H_n^{(2)}$ & $H_n^{(3)}$
0.1	5.7795e-5	6.2487e-5	1.2814e-6
0.2	5.9814e-5	6.425e-5	1.5403e-6
0.3	6.6067e-5	7.0156e-5	1.8656e-6
0.4	7.7362e-5	8.5996e-5	2.326e-6
0.5	8.4324e-5	9.289e-5	2.8203e-6
0.6	9.1617e-5	1.0223e-4	3.5217e-6
0.7	1.019e-4	1.1732e-4	4.2191e-6
0.8	1.2276e-4	1.3991e-4	5.1126e-6
0.9	1.4557e-4	1.6979e-4	6.5654e-6
1.0	1.5277e-4	1.7774e-4	8.0044e-6
1.1	1.8591e-4	2.19e-4	9.6501e-6
1.2	2.2548e-4	2.6528e-4	1.2297e-5
1.3	2.1883e-4	2.7033e-4	1.4832e-5
1.4	2.6811e-4	3.2853e-4	1.8847e-5
1.5	2.9392e-4	3.6358e-4	2.3081e-5
1.6	3.921e-4	4.7871e-4	2.7679e-5
1.7	4.3424e-4	5.3959e-4	3.5531e-5
1.8	4.7715e-4	6.1122e-4	4.2525e-5
1.9	6.2531e-4	8.0041e-4	5.5871e-5
2.0	5.9888e-4	7.9842e-4	6.8209e-5

TABLE 4.2.12

THE TABLE BELOW GIVES THE AVERAGE SQUARE OF DIFFERENCES BETWEEN THE THREE ESTIMATORS WHEN 1000 SAMPLES OF 1000 WERE USED

x	$H_n^{(1)}$ & $H_n^{(2)}$	$H_n^{(1)}$ & $H_n^{(3)}$	$H_n^{(2)}$ & $H_n^{(3)}$
0.1	1.8977e-5	1.9841e-5	3.1535e-7
0.2	1.9515e-5	2.0713e-5	3.8187e-7
0.3	2.2403e-5	2.4114e-5	4.6434e-7
0.4	2.6447e-5	2.86e-5	5.7573e-7
0.5	2.8674e-5	3.0723e-5	6.9937e-7
0.6	3.3952e-5	3.6882e-5	8.5135e-7
0.7	3.4862e-5	3.8122e-5	1.0648e-6
0.8	3.7495e-5	4.1933e-5	1.2836e-6
0.9	4.6376e-5	5.1967e-5	1.5913e-6
1.0	4.6097e-5	5.1467e-5	1.9181e-6
1.1	5.8931e-5	6.721e-5	2.4568e-6
1.2	7.1234e-5	8.1327e-5	2.8928e-6
1.3	6.895e-5	7.9883e-5	3.6062e-6
1.4	9.669e-5	1.1368e-4	4.5053e-6
1.5	9.5438e-5	1.1198e-4	5.5015e-6
1.6	1.0726e-4	1.2686e-4	6.6899e-6
1.7	1.371e-4	1.6531e-4	8.5294e-6
1.8	1.449e-4	1.7963e-4	1.0819e-5
1.9	1.6063e-4	1.9739e-4	1.2706e-5
2.0	1.8855e-4	2.3088e-4	1.5792e-5

TABLE 4.2.13

THE TABLE BELOW GIVES THE VALUES OF TOTAL AVERAGE SQUARE ERROR FOR EACH SAMPLE SIZE AND ESTIMATOR

Estimator	n = 100	n = 500	n = 1000
$H_n^{(1)}$	3.71370	1.57973	1.07700
$H_n^{(2)}$	3.41949	1.56222	1.07088
$H_n^{(3)}$	3.25670	1.54730	1.06569

TABLE 4.2.14

THE TABLE BELOW GIVES THE VALUES OF TOTAL AVERAGE SQUARE DIFFERENCE BETWEEN ESTIMATORS FOR EACH SAMPLE SIZE

Estimator	n = 100	n = 500	n = 1000
$H_n^{(1)}$ & $H_n^{(2)}$	0.08814	0.00468	0.00144
$H_n^{(1)}$ & $H_n^{(3)}$	0.01260	0.00576	0.00170
$H_n^{(2)}$ & $H_n^{(3)}$	0.01080	0.00035	8.265e-5

TABLE 4.2.15

THE TABLE BELOW GIVES THE VALUES OF RELATIVE TOTAL SQUARE DIFFERENCE BETWEEN s^2 AND $\hat{\sigma}^2$ FOR EACH SAMPLE SIZE

Estimator	n = 100	n = 500	n = 1000
$H_n^{(1)}$	0.01762	0.00224	0.00067
$H_n^{(2)}$	0.00259	0.00168	0.00063
$H_n^{(3)}$	0.00314	0.00125	0.00060

TABLE 4.2.16

THE TABLE BELOW GIVES THE K-S Z-VALUES AND TWO TAILED P-VALUES FOR THE THREE ESTIMATORS WHEN 1000 SAMPLES OF 100 WERE USED

	$H_n^{(1)}$		$H_n^{(2)}$		$H_n^{(3)}$	
x	K-S Z	2-TAILED P	K-S Z	2-TAILED P	K-S Z	2-TAILED P
0.1	1.122	0.161	2.764	0.000*	3.053	0.000*
0.2	1.266	0.081	1.921	0.001*	2.235	0.000*
0.3	1.263	0.082	1.709	0.006*	2.023	0.001*
0.4	0.832	0.493	1.015	0.254	1.348	0.053
0.5	1.391	0.042*	1.814	0.003*	2.197	0.000*
0.6	1.633	0.010*	1.816	0.003*	2.239	0.000*
0.7	1.387	0.043*	2.084	0.000*	2.398	0.000*
0.8	1.391	0.042*	1.764	0.004*	2.193	0.000*
0.9	1.191	0.117	1.125	0.159	1.594	0.012*
1.0	1.180	0.123	1.480	0.025*	2.053	0.000*
1.1	1.328	0.059	1.714	0.006*	2.364	0.000*
1.2	1.199	0.113	1.446	0.031*	1.963	0.001*
1.3	1.328	0.059	1.784	0.003*	2.429	0.000*
1.4	2.087	0.000*	2.313	0.000*	2.864	0.000*
1.5	1.644	0.009*	2.191	0.000*	2.851	0.000*
1.6	2.498	0.000*	2.796	0.000*	3.377	0.000*
1.7	1.884	0.002*	2.336	0.000*	2.692	0.000*
1.8	2.311	0.000*	2.861	0.000*	3.442	0.000*
1.9	1.887	0.002*	2.507	0.000*	3.405	0.000*
2.0	1.644	0.009*	2.352	0.000*	3.242	0.000*

* - INDICATE THE HYPOTHESIS THAT THE ESTIMATOR IS ASYMPTOTICALLY NORMAL WITH MEAN H(x) AND VARIANCE $a_n H(x)/n(1-F(x))$ IS REJECTED AT 5% SIGNIFICANCE LEVEL.

TABLE 4.2.17

THE TABLE BELOW GIVES THE K-S Z-VALUES AND TWO TAILED P-VALUES FOR THE THREE ESTIMATORS WHEN 1000 SAMPLES OF 500 WERE USED

	$H_n^{(1)}$		$H_n^{(2)}$		$H_n^{(3)}$	
x	K-S Z	2-TAILED P	K-S Z	2-TAILED P	K-S Z	2-TAILED P
0.1	1.365	0.048*	1.235	0.095	1.147	0.144
0.2	0.790	0.560	0.733	0.656	0.647	0.796
0.3	0.923	0.362	0.930	0.352	1.029	0.241
0.4	0.867	0.440	0.859	0.452	0.895	0.400
0.5	1.095	0.181	1.172	0.128	1.264	0.082
0.6	0.547	0.926	0.506	0.960	0.451	0.987
0.7	1.112	0.168	1.381	0.044*	1.505	0.022*
0.8	1.263	0.082	1.429	0.034*	1.548	0.017*
0.9	0.857	0.454	0.926	0.357	0.950	0.327
1.0	0.713	0.690	0.690	0.728	0.787	0.566
1.1	1.069	0.203	1.349	0.053	1.509	0.021*
1.2	1.071	0.201	0.948	0.329	0.853	0.461
1.3	0.896	0.398	1.108	0.172	1.250	0.088
1.4	1.202	0.111	1.089	0.187	0.973	0.301
1.5	0.760	0.610	0.778	0.581	0.925	0.360
1.6	1.040	0.230	1.185	0.120	1.334	0.057
1.7	0.713	0.689	0.915	0.372	1.102	0.176
1.8	1.050	0.220	1.337	0.056	1.525	0.019*
1.9	1.347	0.053	1.057	0.213	0.931	0.352
2.0	1.151	0.141	1.296	0.070	1.502	0.022*

* - INDICATE THE HYPOTHESIS THAT THE ESTIMATOR IS ASYMPTOTICALLY NORMAL WITH MEAN $H(x)$ AND VARIANCE $\alpha_n H(x)/n(1-F(x))$ IS REJECTED AT 5% SIGNIFICANCE LEVEL.

TABLE 4.2.18

THE TABLE BELOW GIVES THE K-S Z-VALUES AND TWO TAILED P-VALUES FOR THE THREE ESTIMATORS WHEN 1000 SAMPLES OF 1000 WERE USED

	$H_n^{(1)}$		$H_n^{(2)}$		$H_n^{(3)}$	
x	K-S Z	2-TAILED P	K-S Z	2-TAILED P	K-S Z	2-TAILED P
0.1	1.078	0.195	1.049	0.221	1.029	0.240
0.2	0.737	0.649	0.701	0.710	0.656	0.783
0.3	0.594	0.872	0.631	0.820	0.659	0.779
0.4	0.810	0.528	0.901	0.392	0.848	0.468
0.5	1.072	0.200	1.030	0.240	1.096	0.181
0.6	1.240	0.092	1.261	0.083	1.327	0.059
0.7	0.625	0.830	0.612	0.847	0.599	0.865
0.8	1.250	0.088	1.348	0.053	1.448	0.030*
0.9	0.956	0.320	1.094	0.178	1.185	0.120
1.0	1.063	0.208	1.058	0.213	1.130	0.155
1.1	1.328	0.059	1.418	0.036*	1.345	0.054
1.2	1.260	0.084	1.340	0.055	1.408	0.038*
1.3	0.795	0.552	0.840	0.481	0.904	0.387
1.4	0.780	0.577	0.751	0.626	0.649	0.793
1.5	0.940	0.340	1.165	0.132	1.277	0.077
1.6	0.948	0.330	0.988	0.283	1.105	0.174
1.7	0.769	0.595	0.793	0.556	0.912	0.376
1.8	1.228	0.098	1.026	0.243	0.934	0.347
1.9	0.802	0.540	0.784	0.571	0.883	0.417
2.0	1.179	0.124	1.229	0.097	1.349	0.053

* - INDICATE THE HYPOTHESIS THAT THE ESTIMATOR IS ASYMPTOTICALLY NORMAL WITH MEAN H(x) AND VARIANCE $a_n H(x)/n(1-F(x))$ IS REJECTED AT 5% SIGNIFICANCE LEVEL.

CHAPTER V
CENSORED DATA MODEL

The censored data model is one of the major models in the survival analysis. Its use ranges from clinical studies of patient survival time, reliability studies of different mechanisms to studies of some geological features. The classical example is a medical follow–up study (say a cancer treatment study over a fixed time or open ended period.) In many cases we cannot observe the true random survival times $X_1,...,X_n$ of the n patients due to some patients dropping out of the study, live withdrawal of patients from accidents, the study being terminated, etc. . We then say the random variable X_i has been censored by another random variable Y_i. On the other hand, we observe only the random vectors $(d_1,Z_1),...,(d_n,Z_n)$ where

$$d_i = I_{[Z_i=X_i]} \quad \text{and } Z_i = \min\{X_i,Y_i\}, \qquad i = 1,...,n.$$

If Y_i has support (o,∞), the observations are right censored. There has been extensive literature on the hazard rate of the i.i.d. random variables $X_1,...,X_n$ with common distribution F whenever $Y_1,...,Y_n$ are also i.i.d. random variables with common distribution G. Here we define the degree of censorship β as in Koziol and Green (1976)'s model:

$$1 - G(x) = [1 - F(x)]^{\beta}. \tag{5.1.1}$$

(Here $\beta = o$ corresponds to no censoring). For example, if F and G are exponential distributions with parameters λ and μ correspondingly, then

$$1 - G(x) = e^{-\mu x}$$

$$= [e^{-\lambda x}]^{\frac{\mu}{\lambda}}$$

$$= [1 - F(x)]^{\frac{\mu}{\lambda}}.$$

That is, the degree of censorship is μ/λ.

5.1 APPLICATION OF THE ESTIMATORS

Now suppose f and g are the densities of the distribution F and G, then

$$P[Y_i > u, X_i = u] = P[X_i = u, d_i = 1]$$

$$\approx [1 - G(u)]f(u).$$

Let

$$\tilde{F}(x) = P(Z_i \leq x, d_i = 1],$$

then

$$\tilde{F}(x) = \int_{-\infty}^{x} [1 - G(u)]f(u)du$$

and the density

$$\tilde{f}(x) = [1 - G(x)]f(x).$$

Suppose H is the common distribution function of Z_i's, then

$$1 - H(x) = P[Z_i > x]$$

$$= P[X_i > x \text{ and } Y_i > x]$$

$$= [1 - F(x)][1 - G(x)].$$

Therefore,

$$\frac{\tilde{f}(x)}{1-H(X)} = \frac{[1-G(x)]f(x)}{[1-G(x)][1-F(x)]}$$

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

$$= h(x).$$

To provide an estimate of the hazard rate based on censored data, we need to obtain an estimator for $\tilde{f}(x)/(1 - H(x))$. One simple method is the 'reduced sample' estimate by considering only those patients who died of cancer under study. That is, we only consider the random variable $Z_i = X_i$. Now denote this reduced random sample as $(R_1,R_2,\ldots,R_{n'})$ where

$n' = \sum_{i=1}^{n} d_i$ is the reduced sample size. Let

$$R_{n'}(x) = \frac{1}{n'}(\text{number of } R_i\text{'s} \leq x),$$

then the three modified estimators are as follows:

$$h_{n'}^{(1)}(x) = \frac{\frac{1}{n} \sum_{i=1}^{n'} \delta_{n'}(X-R_i)}{1 - R_{n'}(x)},$$

$$h_{n'}^{(2)}(x) = \frac{n'}{n} \sum_{i=1}^{n'} \delta_{n'}(x - R^{(i)}) \frac{1}{n'-i+1},$$

$$h_{n'}^{(3)}(x) = \frac{n'}{n} \sum_{i=1}^{n'} \delta_{n'}(x - R^{(i)}) \log \left[1 + \frac{1}{n'-i+1} \right].$$

Where the $R^{(i)}$'s are the order statistics. Intuitively, the asymptotic equivalence still holds here as the estimators are almost the same except for multiplication by the factor n'/n where $n'/n \rightarrow 1/(1 + \beta)$ as $n \rightarrow \infty$. However, if the degree of censorship increases, the reduced sample size decreases. That is censorship strongly affects the estimators under this reduced sample technique and a poor estimation will be obtained if the degree of the censorship is high. A better estimator has been discussed in Blum and Susarla (1980). However, I do not introduce it here as it is beyond the scope of studying the original estimators.

5.2 NUMERICAL EXAMPLES

In studying the effect of censorship on the reduced sample method, I drew some reduced samples from populations with smooth hazard functions. The procedures were those used in Chapter 4. I first used IMSL to generate single samples of 1000 observations with: (i) X exponentially distributed, mean $\lambda = 1$; (ii) Y exponentially distributed, mean $\mu = 0.1$. I then set $Z = \min(X,Y)$ and sorted out all the $Z_i = X_i$ to produce the reduced sample. This resulted in a degree of censorship of $\beta = 0.1$. The entire procedure was repeated for $\beta = 0.5$ and 1 ($\mu = 0.5$ and 1). To avoid rewriting the programme, I used (3.2.2) $\delta_{n'} = \sqrt{n'}w\,(\sqrt{n'}x)$ as in Chapter 4, and computed $h_{n'}^{(1)}(x)$, $h_{n'}^{(2)}(x)$ and $h_{n'}^{(3)}(x)$ with $x = 0.1, 0.2, \ldots, 2.0$. The results are found in Tables 5.2.1, 5.2.2 and 5.2.3 respectively.

To compare results, I also computed the sum of square error i.e. $\sum_{x=0.1}^{2.0} [h_{n'}^{(i)}(x) - h(x)]^2$. These values are provided in Table 5.2.4.

TABLE 5.2.1

THE TABLE BELOW GIVES THE VALUES OF $H_n^{(1)}(x)$
UNDER DIFFERENT DEGREE OF CENSORSHIP

x	β=0, n=1000	β=0.1, n'=917	β=0.5, n'691	β=1, n'=507
0.1	1.04711	1.0797	1.18068	1.08983
0.2	1.20411	1.17794	1.24013	1.15889
0.3	1.00877	1.06301	1.03978	1.3239
0.4	1.06566	1.06802	1.13843	1.22512
0.5	0.88147	0.86428	0.84688	1.01553
0.6	1.08421	1.00921	1.09383	1.05253
0.7	0.7608	0.78078	1.04348	0.90126
0.8	0.83028	0.89817	0.97944	0.95133
0.9	1.01757	0.98007	0.8073	0.77516
1.0	1.00459	1.15318	1.12784	0.92335
1.1	0.81207	0.80198	0.87458	0.65501
1.2	0.4681	0.54879	0.61056	0.68496
1.3	0.96342	0.96587	0.96994	1.06195
1.4	0.7187	0.63983	0.57971	0.43908
1.5	0.7187	0.63754	0.86497	1.17517
1.6	1.31762	1.19692	1.03796	1.59823
1.7	0.9154	0.99822	0.86497	0
1.8	1.13615	1.13976	1.60273	0.60084
1.9	0.81607	0.89576	0.20183	1.07025
2.0	0.43921	0.61435	1.08121	0

TABLE 5.2.2

THE TABLE BELOW GIVES THE VALUES OF $H_n^{(2)}(x)$
UNDER DIFFERENT DEGREE OF CENSORSHIP

x	β=0, n=1000	β=0.1, n′=917	β=0.5, n′691	β=1, n′=507
0.1	1.04923	1.08355	1.19735	1.1051
0.2	1.20169	1.17375	1.23337	1.14199
0.3	1.00702	1.06022	1.03848	1.3248
0.4	1.05777	1.06095	1.12113	1.21493
0.5	0.8794	0.8611	0.83939	0.98189
0.6	1.08567	1.00965	1.09917	1.07469
0.7	0.76414	0.7829	1.0491	0.89917
0.8	0.82482	0.89247	0.97323	0.94857
0.9	1.02171	0.97904	0.81008	0.76685
1.0	0.99394	1.13869	1.10752	0.91189
1.1	0.81224	0.80076	0.8688	0.66667
1.2	0.46968	0.54994	0.60823	0.65303
1.3	0.95685	0.95999	0.95232	1.02889
1.4	0.72848	0.65044	0.583	0.43927
1.5	0.72181	0.63765	0.86048	1.17926
1.6	1.32761	1.20459	1.01723	1.60858
1.7	0.90141	0.98575	0.85877	0
1.8	1.12651	1.13194	1.5758	0.58582
1.9	0.81362	0.89247	0.19744	1.00962
2.0	0.44076	0.61994	1.13669	0

TABLE 5.2.3

THE TABLE BELOW GIVES THE VALUES OF $H_n^{(3)}(x)$
UNDER DIFFERENT DEGREE OF CENSORSHIP

x	β=0, n=1000	β=0.1, n′=917	β=0.5, n′691	β=1, n′=507
0.1	1.04865	1.08289	1.19634	1.10376
0.2	1.20094	1.17293	1.23214	1.14028
0.3	1.00632	1.05939	1.03727	1.32228
0.4	1.05694	1.06001	1.1196	1.21199
0.5	0.87864	0.86024	0.83806	0.97898
0.6	1.08464	1.00854	1.09716	1.07081
0.7	0.76334	0.78194	1.04686	0.89525
0.8	0.82388	0.89128	0.97086	0.94366
0.9	1.02044	0.9776	0.80783	0.76219
1.0	0.99258	1.13683	1.10398	0.90531
1.1	0.81102	0.79932	0.86562	0.66116
1.2	0.46891	0.54886	0.6057	0.64688
1.3	0.95515	0.95792	0.94781	1.01744
1.4	0.72709	0.64892	0.57991	0.43372
1.5	0.72031	0.63603	0.85542	1.16214
1.6	1.32452	1.20112	1.01017	1.57685
1.7	0.89908	0.98258	0.85207	0
1.8	1.12317	1.12776	1.56076	0.57127
1.9	0.81102	0.8889	0.19532	0.98092
2.0	0.43923	0.61718	1.12267	0

TABLE 5.2.4

THE TABLE BELOW GIVES THE VALUES OF SUM OF SQUARE ERROR

Estimator	β=0, n=1000	β=0.1, n′=917	β=0.5, n′=691	β=1 n′=507
$H_{n'}^{(1)}$	1.10849	0.87038	1.58890	3.35003
$H_{n'}^{(2)}$	1.10521	0.85508	1.57684	3.38463
$H_{n'}^{(3)}$	1.10875	0.85938	1.56721	3.36806

As indicated in Table 5.2.4, the total square error for $\beta = 1$ is extremely large compared with no censorship for all three estimators. This was expected since the sample size was reduced by almost one–half when $\beta = 1$. I did not analyze these simulation further as they would resemble the results of Chapter 4.

REFERENCES

1. Ahmad, I.A. (1976). Uniform strong convergence of the generalized failure rate estimate. Bull. Math. Statist. 17, 77–84.

2. Ahmad, I. and Lin, Pi–Erh (1977). Nonparametric estimation of a vector–valued bivariate failure rate. Ann. Statist. 5, 1027–1038.

3. Bickel, P.J. and Rosenblatt, M. (1973). On some global measures of the deviations of density function estimates. Ann. Stat. 1, 1071–1095.

4. Billingsley, P. (1986). Probability and Measure, Second Edition. Wiley, New York.

5. Blum, J.R. and Susarla, V. (1980). Maximal deviation theory of density and failure rate function estimates based on censored data. P.R. Krishnaiah, ed., Multivariate Analysis Vol. V (North–Holland, Amsterdam, 1980), 213–222.

6. Cox, D.R. and Lewis, P.A. (1966). The Statistical Analysis of Series of Events. Methuen, London.

7. Csörgö, M. and Rëvësz, P. (1981). Strong Approximations in Probability and Statistics. Academic Press, New York.

8. Gaisky, V.N. (1966). On the time distribution of large, deep earthquakes from the Pamir–Hindu–Kush. Physics of solid Earth 6 (English translation), 343–349.

9. Grenander, U. (1956). Estimation of mortality measurement, I and II. Skand. Actuar. Tidskr. 39, 90–96, 125–153.

10. Kimball, A.W. (1960). Estimation of mortality intensities in animal experiments. Biometrics 16, 505–521.

11. Komlós, J. and Major, P. and Tusnády, G. (1975). An approximation of partial sums of independent v.r.'s and the sample d.f.. I.Z. Wahrscheinlichkeitstheorie verw. Geb. 32, 111–131.

12. Koziol, J.A. and Green, S.B. (1976). A Cramér–van Mises statistic for randomly censored data. Biometrika 63, 465–474.

13. Loève, M. (1960). Probability Theory. D. Van Nostrand Company, Princeton.

14. Miller, D.R. and Singpurwalla, N.D. (1978). Large sample estimates and uniform confidence bounds for the failure rate function based on a naive estimator. Tech. Report No. M472, Florida State Univ. Dept. Statist., Tallahassee.

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

16. Rice, J. and Rosenblatt, M. (1976). Estimation of the log survivor function and hazard function. Sankhya A 38, 60–78.

17. Sethuraman, J. and Singpurwalla, N.D. (1981). Large sample estimates and uniform confidence bounds for the failure rate function based on a naive estimator. Ann. Statist. 9, 628–632.

18. Shaked, M. (1978). An estimator for the generalized failure rate function. Tech. Report No. 343, Univ. New Mexico.

19. Udias, A. and Rice, J. (1975). Statistical analysis of microearthquake activity near San Andreas Geophysical Observatory, Hollister, California. Bull. Seismological Soc. Amer. 65, 809–828.

20. Watson, G.S. and Leadbetter, M.R. (1963). On the estimation of the probability density, I. Ann. Math. Stat. 34, 480–491.

21. Watson, G.S. and Leadbetter, M.R. (1964a). Hazard analysis I. Biometrika 51, 175–184.

22. Watson, G.S. and Leadbetter, M.R. (1964b). Hazard analysis II. Sankhyā A 26, 101–116.