This is to certify that the

dissertation entitled

SEQUENTIAL ESTIMATION OF FUNCTIONALS OF THE SURVIVAL CURVE UNDER RANDOM CENSORSHIP WITH APPLICATIONS IN M-ESTIMATION

presented by

MOHAMMAD HOSSEIN RAHBAR

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Statistics and Probability

Joseph C. Gardiner
Major professor

Date July 25, 1988

SEQUENTIAL ESTIMATION OF FUNCTIONALS OF THE SURVIVAL CURVE UNDER RANDOM CENSORSHIP WITH APPLICATIONS IN M-ESTIMATION

by

MOHAMMAD HOSSEIN RAHBAR

A DISSERTATION

Submitted to

MICHIGAN STATE UNIVERSITY

in partial fulfilment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

Department of Statistics and Probability

1988

ABSTRACT

Sequential estimation of functionals of the survival curve under random censorship with applications in M-estimation.

by

Mohammad Hossein Rahbar

Sequential point and interval estimation procedures for functionals of the survival curve F (of the form $\int \psi dF$ and $\int F d\psi$) are considered when the underlying observations may be subject to random censorship.

In the point estimation problem, the loss is measured by the sum of the squared error of the estimator and cost of observations made with per unit cost c being constant.

The sequential estimator defined here is shown to be risk efficient and normal as c tends to zero under certain regularity conditions on functions ψ, F and the censoring distribution.

For the interval estimation, the sequential procedure is shown to be consistent and the corresponding stopping rule is shown to be efficient as the width of the interval decreases to zero.

In both estimation problems, the asymptotic distribution of the stopping rule is obtained.

Finally, as an application the consistency and efficiency of a sequential fixed width interval estimation procedure using M-estimation is shown for the location parameter in a location model when the error distribution is symmetric and continuous and the censoring distribution is continuous but unknown.

To my parents and my wife and daughter

ACKNOWLEDGEMENTS

I wish to express my sincere thanks to Professor Joseph C. Gardiner for his excellent guidance, continuous encouragement, careful reading of this thesis and especially his patience during the preparation of this dissertation.

I would also like to thank Professors James Hannan, Hira L. Koul, Habib Salehi and R.V. Ramamoorthi for serving on my guidance committee. My special thanks go to Professors James Hannan and Hira Koul for their helpful discussions and careful reading of this dissertation and for giving useful suggestions, which helped greatly in the final presentation of this thesis.

I would also like to thank Professor James Stapleton for his help in doing some simmulation studies.

I would like to thank my wife Afsaneh, my daughter Elaheh and my parents for their moral support during the preparation of this dissertation.

Finally, my special thanks go to secretaries Cathy Sparks, Loretta Ferguson and JoAnn Peterson for their help in typing the manuscript.

TABLE OF CONTENTS

Chapter 0

Introduction and summary:

The problem of estimation of various functionals of the survival curve from censored data is of fundamental importance in epidemiological and reliability studies, clinical trials and life testing. In most clinical trials ethical reasons and the paucity of laboratory specimens compel consideration of statistical procedures in which the sample size is not specified in advance of experimentation.

In this thesis, sequential point and interval estimation procedures for functionals of the survival curve F (of the form $\int\psi dF$ and $\int F d\psi$) are considered when the underlying observations may be subject to random censorship. When censoring is present, as is generally the case in several survival studies, one observes a random sample $\{(Z_i,\delta_i)\colon 1\le i\le n\}$ where $Z_i= \min(X_i,Y_i)$ and $\delta_i= [X_i\le Y_i]$ with life times X_i's, having common survival curve F and censoring times Y_i's, independent of X_i's, having common survival curve G and [A] denoting the indicator function of the set A.

Consider the natural estimator of θ, $\hat{\theta}_n=\int\psi d\hat{F}_n$, where $\hat{F}_n$ is a suitable estimator of F based on the above random sample.

In the point estimation problem we consider the loss structure $L_n(c) = a(\hat{\theta}_n - \theta)^2 + cn$ where a is a given positive constant. The associated risk is $R_n(c)=E(L_n(c))$. The problem is to determine the sample size which minimizes the risk $R_n(c)$ for a given positive cost per observation c.

In the interval estimation problem we construct a confidence interval for $\theta(F)$ of prescribed width 2d and coverage probability $(1-2\alpha)$ where $0<2\alpha<1$. In either case, the best fixed sample size procedure (BFSSP), say n_0, possessing the desired property depends on unknown functions F, G and therefore no fixed sample size procedure minimizes the risk universally for all F, G. In order to resolve this difficulty we define a sequential sampling rule (often called stopping time) similar to the rule considered by Robbins (1959). Essentially, this rule, say N, samples observations sequentially, updates a suitable estimator of F, G in the BFSSP and stops sampling as soon as the number of observations exceeds that of the estimated BFSSP. Thus one is led to solve these problems using sequential procedures.

In the point estimation problem, the performance of the procedure $(N,\hat{\theta}_N)$ is usually measured by (i) Relative risk, (R^*/R_0), and (ii) The regret, $r^*(c) = R^* - R_0$ where R^* is the risk due to using stopping time N and R_0 is the risk in the BFSSP. In (i) if $\lim (R^*/R_0) = 1$, as c tends to zero, the sequential procedure is said to be risk efficient. In (ii) if $r^*(c) = O(c)$, the sequential procedure is said to have bounded regret.

In the interval estimation problem, the performance of (N,I_N) is measured by the (iii) Coverage probability, $P[\theta\epsilon I_N]$ and (iv) Expected sample size, E(N), where I_N is the fixed width interval estimator using sample size N. In (iii) if $\lim P[\theta\epsilon I_N]=1-2\alpha$, the procedure is said to be consistent. In (iv) if $EN/n_0 \to 1$, the procedure is said to be efficient where in (iii) and (iv) limits are taken as d tends to zero.

We now give a brief review of the literature on sequential point and interval estimation.

Sequential point and interval estimation problems have received considerable attention ever since the fundamental paper of Robbins (1959) for the estimation of the mean of a normal population. He considered $X_1,X_2,\ldots,X_n$ iid $N(\mu,\sigma^2)$ and the loss function $L_n = a|\overline{X}_n - \mu| + n$, where $\overline{X}_n$ is the sample mean based on n observations and a is some positive known

constant. He assumed that the cost of sampling is proportional to the sample size and showed that when σ is known the BFSSP is $n_0=(a\sigma/\sqrt{2\Pi})^{2/3}$. For this sample size the risk is $R_0=E(L_{n_0})=3n_0$. All this presupposes that we know σ. When σ is unknown he suggested to take the sample size $N = \inf\{n\geq 3: n\geq(a\hat{S}_n/\sqrt{2\Pi})^{2/3}\}$ where $\hat{S}_n^2$ is the sample variance. Starr (1966) proved the risk efficiency and Starr and Woodroofe (1969) showed that the regret is bounded in the above case.

In the nonparametric context, Ghosh and Mukhopadhyay (1979) were first to prove the risk efficiency for the mean problem under the condition that the eighth moment is finite. Sen and Ghosh (1981) considered sequential point estimation of estimable parameters based on U–statistics under the condition that $E|g|^{2+s}< \infty$, for some positive real number s, where g is the kernel corresponding to the parameter. Estimation of the mean is a particular case with g as the identity function. Chow and Yu (1981) have proved risk efficiency for the mean problem provided an rth moment is finite for some $r>2$. Sequential point estimation of location based on some R–, L–, and M–estimators are discussed in Sen (1980) and Jureckova and Sen (1981). Sen's book (1981) has an excellent survey of the above mentioned articles.

Gardiner and Susarla (1983) were the first to consider the sequential point estimation of the mean problem, in a nonparametric context when censorship is present. They did not find the asymptotic distribution of the stopping time except for the case of an exponential survival time. (See Gardiner, Susarla and van Ryzin (1985b)).

In this thesis we propose a sequential point and an interval estimation procedures for functionals of the form $\int\psi dF$ and $\int F d\psi$ when the underlying observations may be subject to random censorship. We shall show that the sequential point estimation procedure is risk efficient and asymptotically normal as c tends to zero, under certain regularity conditions on functions ψ, F and the censoring distribution. For the interval estimation problem, the sequential procedure is shown to be consistent and efficient as the width of the confidence interval decreases to zero. In both estimation problems, the asymptotic distribution of the underlying stopping time is obtained. Thus our results are generalizations of Gardiner and Susarla (1983) and Gardiner, Susarla and van Ryzin (1985b).

In Chapter 1, we collect various preliminaries and necessary prerequisities of asymptotic properties of the product–limit (P–L) estimator. Most of the results are taken from Gill (1983), Földes and Rejtõ (1981), Cheng (1984), Lo and Singh (1986), Gardiner, Susarla and van Ryzin (1985a) and Schick, Susarla and Koul (1987). Hence some of the proofs have been omitted. We also state the results regarding to the asymptotic normality and the almost sure representation of the estimator of σ^2, the asymptotic variance of $n^{1/2}(\int\psi d\hat{F}_n-\int\psi dF)$, which are new and crucial for obtaining the asymptotic distribution of stopping times in Chapters 2 & 3 and we give an almost sure representation of $n^{1/2}(\int\psi d\hat{F}_n-\int\psi dF)$, for ψ in a class of functions Ψ_1 (defined later in Chapter 1).

Chapter 2 is divided into three sections. The first two develop our model and some examples are discussed. In Section 3 the risk efficiency of the sequential point estimation procedure and the asymptotic normality of the underlying stopping time are presented.

In Chapter 3 we discuss the properties of the sequential fixed width confidence interval and the related theorems.

Chapter 4 deals with a sequential fixed width confidence interval procedure for the location parameter of a location model under random censorship when the distribution of the error is symmetric around zero but unknown.

Finally, proofs of Theorems 1.3 and 1.4, the asymptotic normality, almost sure representation and a rate of convergence of the estimator of σ^2, the asymptotic variance of $n^{1/2}(\int\psi d\hat{F}_n - \int\psi dF)$, which are crucial for obtaining the asymptotic distribution of the stopping time in Chapters 2 & 3, are placed in Appendices.

Chapter 1

1.1 Preliminaries and some notations

Let $\{X_i: i\geq 1\}$ be a sequence of nonnegative iid rv's (survival times) with continuous survival function F on $\mathbf{R}^+=[0,\infty)$, $F(0)=1$. The corresponding censoring rv's $\{Y_i: i\geq 1\}$ are also assumed iid, independent of $\{X_i: i\geq 1\}$, with continuous survival function G taking values on $\mathbf{R}^+$ and $G(0)=1$. The observable data are $\{(Z_i,\delta_i): i\geq 1\}$ where $Z_i= \min(X_i,Y_i)$, $\delta_i= [X_i\leq Y_i]$ and $[A]$ denotes the indicator function of the set A. For an estimator of F we select the product–limit (P–L) estimator, $\hat{F}_n$, which was first introduced by Kaplan and Meier (1958), based on $\{(Z_i,\delta_i): 1\leq i\leq n\}$ defined by

$$(1.1.1)\qquad \begin{aligned} \hat{F}_n(t) &= \prod_{i=1}^{n} \left[\frac{K(Z_i)-1}{K(Z_i)}\right]^{[Z_i\leq t,\delta_i=1]} && \text{for } t<Z_{(n)} \\ &= 0 && \text{for } t\geq Z_{(n)} \end{aligned}$$

where $K(t)= 1 + \sum_{i=1}^{n} [Z_i>t]$, $t\geq 0$, and $Z_{(1)}< Z_{(2)}< \ldots.< Z_{(n)}$ are the order statistics of $\{Z_i: 1\leq i\leq n\}$. By the continuity of F and G ties among the observations may be disregarded with probability one. Throughout we shall need the following notations.

Let (X_j,B_j), $j\geq 1$ be copies of the $\mathbf{R}^+\times \{0,1\}$ with Borel σ–fields and $(X^*,A^*) = \prod_{i=1}^{\infty}(X_i,B_i)$. Let $P = P_{F,G}$ denote the product measure induced by $\{(Z_i,\delta_i): i\geq 1\}$ on A^* and $E = E_{F,G}$ denote the expectation under P.

Let $\mathbf{P} = \{P: F, G \in \mathbf{F}\}$ where $\mathbf{F}$ = {the class of continuous survival functions} and $\theta = \theta(F)$ be a real valued functional on $\mathbf{F}$. The estimator $\hat{\theta}_n= \theta(\hat{F}_n)$ of θ has been studied by many authors including Breslow and Crowley (1974), Wellner (1982), Gardiner and Susarla (1983), Gill (1983), and Millar (1985). For $\{\hat{\theta}_n\}$ to be consistent for θ, certain conditions have to be satisfied by the pair (F,G). For example, in Schick, Susarla and Koul (1987) is stated that "when estimating $\theta(F)$, the p-th quantile of (1–F), we need at least the condition $G(\theta(F))>0$, for the sample quantile to be consistent. Generally, $\theta(\hat{F}_n)$ is not consistent for $\theta(F)$ if $\theta(F)$ depends on parts of F that lie beyond the upper support point $\tau_G = \inf\{x: G(x)=0\}$ of G." Similarly define τ_F and let $\tau=\tau_H=\min(\tau_F,\tau_G)$. Let (Z,δ) be a copy of (Z_1,δ_1). Let $\approx{H}(t) = P[Z\le t,\delta=1]$, $\tilde{H}(t) = P[Z\le t,\delta=0]$ and $H(t) = P[Z>t]$. Note that $\tilde{H}(s) = -\int_0^s FdG$, $\approx{H}(s) = -\int_0^s GdF$ and $H(t) = 1 - \tilde{H}(t) - \approx{H}(t)$.

Throughout, except in Chapter 4, all unspecified integrals are considered on R^+. Throughout this thesis L^{-r} stands for $(1/L)^r$ for any positive function L; $\xrightarrow{D}$ denotes "convergence in distribution"; a.s. stands for almost sure with respect to probabitity measure P; $N(\mu,\sigma^2)$ will stand for the normal distribution with mean μ and variance σ^2 and the index in the summations runs from 1 through n unless it is otherwise specified; Ψ denotes the class of all real valued monotone functions on R^+ and let

$$\Psi_1 = \{\psi\in\Psi : \psi \text{ is constant on } [T,\infty) \text{ for some } T<\tau\},$$

$$A(t) = \int_{[t,\infty)} F d\psi,\quad A_n(t) = \int_{[t,\infty)} \hat{F}_n d\psi, t\ge 0,$$

$$C(t) = \int_0^t H^{-2} d\approx{H},\quad 0\le t<\tau,$$

$$\sigma^2 = \int A^2 dC\ ,\quad \hat{\sigma}_n^2 = \int A_n^2\, n^2\, K^{-2} d\approx{H}_n,$$

(1.1.2) $$\Gamma_{ij}(t) = \int_0^t A^i\, dC^j,\ t\ge 0, \text{ for } i,j = 1,2,3,4,$$

(1.1.3) $$\xi(Z,\delta;t) = C(Z\wedge t) - \{\delta\, H^{-1}(Z)\}[Z\le t],\ t\ge 0$$

and

$$J(Z,\delta) = -\int \xi(Z,\delta;t)F(t)d\psi(t)$$

(1.1.4) $$= \delta A(Z)H^{-1}(Z) - \Gamma_{11}(Z).$$

The following lemma is taken from Theorem 1 of Lo and Singh (1986) and Theorem 3.4 of Gardiner, Susarla and van Ryzin (1985a) and will be stated without proof.

Lemma 1.1: If F and G are continuous and $T<\tau$, then on $[0,T]$ and for $p>0$,

(1.1.5) $$\hat{F}_n(t)- F(t) = F(t)\{n^{-1}\Sigma\ \xi(Z_i,\delta_i;t) + r_n(t)\}$$

with

(1.1.6) $$\sup|r_n(t)| = O((n^{-1}\ln n)^{3/4}) \quad \text{a.s.},$$

and

(1.1.7) $$\sup \|r_n(t)\|_p = O(n^{-1})$$

where $\| \cdot \|_p$ denotes the L_p norm and sup is taken over $[0,T]$ and $\xi(Z,\delta;t)$ is as in (1.1.3).

Assumption 1.1: Let T be a positive constant such that $T<\tau$. Let ψ be a monotone nondecreasing function defined on $\mathbf{R}^+$ such that $\psi(x) = b$, for $x\geq T$, where b is a constant.

Theorem 1.1: Under Assumption 1.1 and assumptions of Lemma 1.1,

(1.1.8) $$n^{1/2}(\int\psi d\hat{F}_n- \int\psi dF) - n^{-1/2}\Sigma J(Z_i,\delta_i) = O(n^{-1/4}\{\ln n\}^{3/4}), \text{ a.s..}$$

Proof: Integration by parts and Assumption 1.1 and (1.1.5) allow us to write

$$\begin{aligned} n^{1/2}(\hat{\theta}_n- \theta) &= n^{1/2}(\int\psi d\hat{F}_n- \int\psi dF) \\ &= - n^{1/2}\int \left(\frac{\hat{F}_n- F}{F}\right) F\, d\psi \\ &= n^{-1/2}\ \Sigma\ J(Z_i,\delta_i) - n^{1/2}r_n^* \end{aligned}$$

where J is as in (1.1.4), $r_n(t)$ is as in (1.1.5) and $r_n^* = \int r_n F d\psi$. Note that by (1.1.6) and Assumption 1.1,

$$\begin{aligned} r_n^* = \int r_n(t)F(t)d\psi(t) &\leq \{\sup_{0\leq t\leq T}|r_n(t)|\} \int F d\psi \\ &= O\ (n^{-1}\ln n)^{3/4} \quad \text{a.s.,} \end{aligned}$$

from which (1.1.8) is immediate. □

Corollary 1.1: Under assumptions of Theorem 1.1,

$$n^{1/2}(\int \psi d\hat{F}_n - \int \psi dF) \xrightarrow{D} N(0,\sigma^2). \tag{1.1.9}$$

Proof: One can show that $J(Z,\delta)$ is a bounded random variable with mean zero and variance σ^2. Therefore by the central limit theorem (CLT) and (1.1.8), (1.1.9) is immediate. □

Theorem 1.2: Under the assumptions of Theorem 1.1 and for $s>0$,

$$\{n^{s/2}|\hat{\theta}_n - \theta|^s\colon n\geq 1\} \text{ is uniformly integrable (UI).}$$

Proof: Recall the representation

$$n^{1/2}(\hat{\theta}_n - \theta) = n^{-1/2}\Sigma J(Z_i,\delta_i) - n^{1/2}r_n^*.$$

Note that it suffices to show that $\{|n^{-1/2}\Sigma J(Z_i,\delta_i)|^s\colon n\geq 1\}$ and $\{|n^{1/2}r_n^*|^s\colon n\geq 1\}$ are both UI. Since for $t>T$, $A(t) = 0$ and $H(T)>0$, $\{J(Z_i,\delta_i)\colon i\geq 1\}$ is a sequence of bounded random variables. Thus an application of the Marcinkiewicz–Zygmund inequality implies that for any $p>0$,

$$\sup_{n\geq 1} E|n^{-1/2}\Sigma J(Z_i,\delta_i)|^p<\infty,$$

which implies $\{|n^{-1/2}\Sigma J(Z_i,\delta_i)|^s\colon n\geq 1\}$ is UI. Recall that $r_n^* = \int r_n(t)F(t)d\psi(t)$. By the Holder inequality, Fubini Theorem and (1.1.7) we have that for any $p>0$,

$$E|r_n^*|^p = O(n^{-p/2}). \tag{1.1.10}$$

Note that (1.1.10) implies that $\{|n^{1/2}r_n^*|^s\colon n\geq 1\}$ is UI, which completes the proof of the theorem. □

Remark 1.1: Schick, Susarla and Koul (1987) gave a sufficient condition for an iid representation of $n^{1/2}(\int \psi d\hat{F}_n - \int \psi dF)$, where $\psi\in\Psi$, the class of all real valued monotone functions on $\mathbf{R}^+$. They have shown that

$$n^{1/2}(\int \psi d\hat{F}_n - \int \psi dF) - n^{-1/2}\Sigma J(Z_i,\delta_i) = o_p(1).$$

Our Theorem 1.1 gives the corresponding almost sure representation for $\psi\in\Psi_1$, a subclass of Ψ.

Theorem 1.3 **(The almost sure representation and asymptotic normality of $\hat{\sigma}_n^2$)**: Let $\psi\in\Psi_1$. For F, G$\in$**F**,

$$n^{1/2}(\hat{\sigma}_n^2 - \sigma^2) = n^{-1/2}\Sigma V_i + n^{1/2}R_{n,5}$$

where

$$V_i = 2W_{i,1} + W_{i,2},$$

$$W_{i,1} = \int\{\int_{\cdot}^{\infty} \xi(Z_i,\delta_i;s)F(s)d\psi(s)\}d\Gamma_{11},$$

$$W_{i,2} = A^2(Z_i)H^{-2}(Z_i)\delta_i - 2\int H^{-1}[Z_i>\cdot]d\Gamma_{21} + \sigma^2,$$

and

$$n^{1/2}R_{n,5} \longrightarrow 0, \text{ a.s.}.$$

Furthermore

(1.1.11) $$n^{1/2}(\hat{\sigma}_n^2 - \sigma^2) \xrightarrow{D} N(0,\gamma^*),$$

where

(1.1.12) $$\gamma^* = 6\int\Gamma_{11}^2\, d\Gamma_{21} - 4\int H^{-1}\Gamma_{11}\, d\Gamma_{31} + \int H^{-2}\, d\Gamma_{41} - \sigma^4.$$

Proof: See Appendix A.

Theorem 1.4: Let $\psi\in\Psi_1$. For F, G$\in$**F**, and for each $\epsilon>0$, and all $r<\infty$,

(1.1.13) $$P[|\hat{\sigma}_n^2 - \sigma^2| \geq \epsilon] = O(n^{-r}).$$

Proof: See Appendix B.

Chapter 2

Sequential point estimation of functionals of the survival curve under random censorship

2.1 Introduction:

In this chapter we consider the sequential point estimation of functionals of the form $\int \psi dF$ and $\int F d\psi$ where $\psi \in \Psi_1$ and $F \in \mathbf{F}$.

Given a sample of size n, $\{(Z_i, \delta_i): 1 \leq i \leq n\}$, we estimate θ by, $\hat{\theta}_n = \int \psi d\hat{F}_n$, subject to the loss function

(2.1.1) $$L_n(c) = a(\hat{\theta}_n - \theta)^2 + cn$$

where a is a positive constant and c is the cost per unit observation. The object is to minimize the risk in estimation by choosing an appropriate sample size. From Corollary 1.1, we have that

(2.1.2) $$n^{1/2}(\hat{\theta}_n - \theta) \xrightarrow{D} N(0, \sigma^2), \text{ as } n \to \infty.$$

Recall that in Theorem 1.2, it is shown that (under certain conditions) the sequence $\{n^{s/2} |\hat{\theta}_n - \theta|^s: n \geq 1\}$ is UI for s>0. Therefore it follows from (2.1.2) that

(2.1.3) $$E(\hat{\theta}_n - \theta)^2 = n^{-1}\sigma^2 + o(n^{-1}), \text{ as } n \to \infty.$$

Now if σ is known, then the risk

(2.1.4) $$R_n(c) = E\, L_n(c) = n^{-1} a\sigma^2 + cn + o(n^{-1})$$

is approximately minimized by the BFSSP, $n_0 \cong b\sigma$, where $b = (a/c)^{1/2}$, with corresponding minimum risk

(2.1.5) $$R_0 = R_{n_0} \cong 2cn_0.$$

However, since σ is unknown, the BFSSP cannot be used and therefore we describe a sequential procedure for choosing a sample size whose risk will be close to R_0 for small c. Let

(2.1.6) $$N_c = \inf\{n \geq n_{1c}: n \geq b(\hat{\sigma}_n + n^{-h})\}$$

where h is a positive constant to be selected later. Since $N_c \geq b N_c^{-h}$ a.s., we may assume $n_{1c} = \mathrm{int}(b^{1/(1+h)})$ where int(x) denotes the greatest integer $\leq$ x. Then our scheme utilizes the estimator $\hat{\theta}_{N_c}$ of θ with associated risk

$$R^* = R_c^* = EL_{N_c} = a\, E(\hat{\theta}_{N_c} - \theta)^2 + c\, EN_c. \tag{2.1.7}$$

We now consider some examples before we present the main results of this chapter.

2.2 Examples

1. **A form of winsorized mean:** Let $T<\tau$, and

$$\psi(x) = (x \wedge T)[x \geq 0] \tag{2.2.1}$$

and $\theta = -\int \psi dF$, $F \epsilon \mathcal{F}$. Note that $\hat{\theta}_n = -\int \psi d\hat{F}_n = -\Sigma \psi(Z_i)\delta_i d_{in}$, where d_{in} is the jump of the P–L estimator at Z_i when a sample of size n is observed. Since in the absence of censorship the jumps of the P–L estimator reduce to (1/n), $\hat{\theta}_n$ reduces to an estimator of θ where the ordinary empirical process F_n replaces $\hat{F}_n$ above. In this case $\hat{\theta}_n$ turns out to be the average of all the observations which lie in [0,T) and all the observations greater or equal to T replaced by T. When using the P–L estimator to estimate $E(X \wedge T)$, the estimator $\hat{\mu} = \int_0^T \hat{F}_n(t)dt$, $T<\tau$, is used. Susarla and van Ryzin (1979) have generalized the estimator to get the mean by taking $\hat{\mu} = \int_0^{M_n} \hat{F}_n(t)dt$, where $M_n \uparrow \infty$, as $n \rightarrow \infty$, with certain restrictions on $\{M_n\}$. Gill (1983) indicates how M_n can be replaced by $Z_{(n)}$ in the estimator $\hat{\mu} = \int_0^{M_n} \hat{F}_n(t)dt$.

Remark 2.2.1 (taken from Remark 4.1 of Gardiner and Susarla (1983))
If the problem of interest is the estimation of the mean survival time observed on the duration [0,T], where $T<\tau$, that is, $E(X \wedge T) = \int_0^T F(u)\, du$, then by taking ψ as in (2.2.1) we get $\sigma_T^2 = \Gamma_{21}(T)$. One encounters this situation of

estimation of $E(X \wedge T)$, T finite in some decrement models. See Gardiner (1982) and Hoem (1976) and (1987). Analogous remarks hold for the problem of sequential estimation of $E(X|X<T) = T - (1-F(T))^{-1}\int_0^T (1-F(u))du$, with, of course, a different expression for the asymptotic variance, σ^2.

2. **A form of winsorized sample moments:**

Let $k>0$, $\psi(x) = (x\wedge T)^k[x\geq T]$ and $F\in\mathbf{F}$. Consider $\hat{\theta}_n = -\int\psi d\hat{F}_n$ as an estimator of $\theta = -\int\psi dF$. Note that if the mean of F is known, we can estimate the variance of F by taking $\hat{\theta}_n = -\int\psi d\hat{F}_n$ where

$$\psi(x) = \psi(x) = (x\wedge T)^2[x\geq T] - \mu_F^2$$

and μ_F denotes the mean of F.

3. **A form of mean residual life:**

The mean residual life function is defined by

$$\rho(t) = F^{-1}(t)\int_t^\infty F(s)ds = - F^{-1}(t)\int_t^\infty (s-t)dF(s).$$

Our interest is in estimation of $\theta(F) = F^{-1}(t)\int_t^T F(s)\, ds$, for a fixed t, $t<T$. If we know $F(t)$, for example at $t = \text{med}(F)$, $F(t) = 1/2$, then our scheme estimates θ by $\hat{\theta}_n = \int F d\psi$, where $\psi(x) = F^{-1}(t)\{((tVx)\wedge T\}[x\geq 0]\}$ and therefore the asymptotic variance of this estimator is $\sigma^2 = F^{-2}(t)\{\Gamma_{21}(T) - \Gamma_{21}(t)\}$. In the case that $F(t)$ is unknown, one can use $\hat{F}_n(t)$ as an estimator of $F(t)$ with, of course, a different expression for the asymptotic variance.

4. **Kaplan–Meier M–estimator**

This example will be discussed in detail in Chapter 4.

2.3 Main results of this chapter:

In the sequel all limits are taken as $c \downarrow 0$ or $b \uparrow \infty$. We shall drop the subscript c in N_c, n_{1c} and on various entities when there is no possibility of confusion.

Remark 2.3.1 Throughout all the proofs are given for $\theta(F)=\int\psi dF$, where ψ is as in Assumption 1.1. Note that handling the case that $\theta(F)=\int Fd\psi$ where ψ is as in Assumption 1.1, is very similar. In the case that ψ is monotone nonincreasing on $\mathbf{R}^+$ and constant on $[T,\infty)$, we can reduce the problem to the above case by writing $\theta(F)=-\int-\psi\, dF$. Hence all of the results hold for $\psi\in\Psi_1$ and θ of the form $\int\psi dF$ and $\int Fd\psi$ where $F\in\mathbf{F}$.

The following results hold under Assumption 1.1.

Theorem 2.3.1: With N defined in (2.1.6) and for each $P\in\mathbf{P}$,

$$n_0^{-1}N \longrightarrow 1 \quad \text{a.s.,} \tag{2.3.1}$$

and

$$E|n_0^{-1}N - 1| \longrightarrow 0. \tag{2.3.2}$$

Theorem 2.3.2 (Risk efficiency): With N and R^* defined in (2.1.6) and (2.1.7),

$$R^*R_0^{-1} \longrightarrow 1. \tag{2.3.3}$$

Theorem 2.3.3 (Asymptotic distribution of N): Let $h>1/2$ and N as in (2.1.6), then

$$n_0^{1/2}(n_0^{-1}N - 1) \xrightarrow{D} N(0,\gamma^*/(4\sigma^4)) \tag{2.3.4}$$

and

$$N^{1/2} - n_0^{1/2} \xrightarrow{D} N(0,\gamma^*/(16\sigma^4)) \tag{2.3.5}$$

where $\gamma^* = 6\int \Gamma_{11}^2 d\Gamma_{21} - 4\int H^{-1}\Gamma_{11}d\Gamma_{31} + \int H^{-2}d\Gamma_{41} - \sigma^4$.

Proof of Theorem 2.3.1: By definition of N_c, $\lim N_c=\infty$, a.s., also if $0<c_1<c_2$, we have

$$N_{c_1} \geq (a/c_1)^{1/2}(\hat{\sigma}_{N_{c_1}} + N_{c_1}^{-h}) > (a/c_2)^{1/2}(\hat{\sigma}_{N_{c_1}} + N_{c_1}^{-h}).$$

Hence by definition of N_c we obtain $N_{c_1} \geq N_{c_2}$ a.s.. Thus N_c is nondecreasing as $c\downarrow 0$. From (1.1.13) and the Borel–Cantelli lemma, it follows that $\{\hat{\sigma}_n^2\colon n\geq 1\}$ is a strongly consistent estimator of σ^2. Since $N_c\uparrow\infty$, a.s., (2.3.6) $$\hat{\sigma}_N^2 \to \sigma^2 \quad \text{a.s..}$$

Recall that $b = (a/c)^{1/2}$. By definition of N, we can write

$$b\,\hat{\sigma}_N \le b\,(\hat{\sigma}_N + N^{-h}) \le N < b\,(\hat{\sigma}_{N-1} + (N-1)^{-h}) + 1.$$

So that on dividing all sides by n_0 and using (2.3.6) we get

$$n_0^{-1}N \to 1, \text{ a.s..}$$

The next lemma is very similar to Lemma 1 of Gardiner and Susarla (1983) which gives a rate on the tail behavior of the stopping rule N which is crucial for the proof of (2.3.2).

Lemma 2.3.1: For each $0<\epsilon<1$ and for any $r<\infty$,

(2.3.7) $$P[N \le n_0(1-\epsilon)] = O(c^{(r-1)/2(1+h)})$$

and

(2.3.8) $$\sum_{n \ge n_0(1+\epsilon)} P[N \ge n] = O(c^{(r-1)/2}).$$

Proof: Recall that $n_0 \cong b\sigma$ and $n_1 = b^{1/(1+h)}$. Let $n_2 = n_{2c} = \text{int}(n_0(1-\epsilon))$ and $n_3 = n_{3c} = \text{int}(n_0(1+\epsilon))$. By definition of N, $N \ge n_1$, a.s.. For sufficiently small c we have $n_1 < n_2$, and on the set $[N \le n_2]$, $n \ge b\hat{\sigma}_n$, for some $n \in \{n_1,\dots,n_2\}$. Therefore, for small c,

$$\begin{aligned}
P[N \le n_2] &\le P[\hat{\sigma}_n \le b^{-1}n, \text{ for some } n \in \{n_1,\dots,n_2\}] \\
&\le P[\hat{\sigma}_n^2 - \sigma^2 \le b^{-2}n_2^2 - \sigma^2, \text{ for some } n \in \{n_1,\dots,n_2\}] \\
&\le P[\hat{\sigma}_n^2 - \sigma^2 \le \{(1-\epsilon)^2 - 1)\sigma^2, \text{ for some } n \in \{n_1,\dots,n_2\}] \\
&\le \sum_{n=n_1}^{n_2} P[|\hat{\sigma}_n^2 - \sigma^2| \ge \epsilon(2-\epsilon)\sigma^2].
\end{aligned}$$

Thus (1.1.13) and the usual integral approximation for sums yield

$$\begin{aligned}
P[N \le n_2] &\le \text{const.} \int_{n_1}^{n_2} x^{-r}dx \\
&= O(n_1^{-(r-1)}) \\
&= O(c^{(r-1)/2(1+h)}).
\end{aligned}$$

To show (2.3.8), consider $n \geq n_3$. Then on the set $[N>n]$, we have that $k<b(\hat{\sigma}_k + k^{-h})$, for all $k \in \{n_1,\ldots,n\}$. For c sufficiently small and $n \geq n_3$,

$$\begin{aligned} P[N>n] &\leq P[\hat{\sigma}_n > b^{-1}n - n^{-h}] \\ &\leq P[\hat{\sigma}_n - \sigma > b^{-1}n_3 - b^{-1}n_0 - n_3^{-h}] \\ &\leq P[\hat{\sigma}_n - \sigma > (1/2)\ \epsilon\sigma] \\ &\leq P[|\hat{\sigma}_n^2 - \sigma^2| > (1/4)\ \epsilon^2\sigma^2]. \end{aligned}$$

The last relation holds because

$$\begin{aligned} \hat{\sigma}_n^2 - \sigma^2 &= (\hat{\sigma}_n - \sigma)^2 + 2\sigma(\hat{\sigma}_n - \sigma) \\ &\geq (1/4)\ \epsilon^2\sigma^2 + \sigma^2\epsilon \\ &\geq (1/4)\ \epsilon^2\sigma^2. \end{aligned}$$

Therefore using an integral approximation for sums and similar arguments as in the previous case leads to (2.3.8). This completes the proof of Lemma 2.3.1. □

Now we are ready to prove (2.3.2). Let $0<\epsilon<1$, $D = [n_2<N\leq n_3]$ and $\overline{D}$ be the complement of the set D. Write

$$Nn_0^{-1} - 1 = Nn_0^{-1}[N\leq n_2] + (Nn_0^{-1} - 1)[D] + Nn_0^{-1}[N>n_3] - [\overline{D}].$$

Note that $N\leq n_2$, implies that $Nn_0^{-1} \leq 1-\epsilon$. Hence

$$\begin{aligned} E|Nn_0^{-1} - 1| &\leq (1-\epsilon)\ P[N\leq n_2] + \epsilon + n_0^{-1} \sum_{n\geq n_3} P\ [N>n] + P[\overline{D}] \\ &= O(c^{(r-1)/2(1+h)}) + \epsilon + n_0^{-1}O(c^{(r-1)/2}) + o(1) \\ &= o(1), \text{ since } r>1 \text{ and } \epsilon \text{ is arbitrary.} \end{aligned}$$

This completes the proof of Theorem 2.3.1. □

Proof of Theorem 2.3.2 (Risk efficiency): Note that

$$\begin{aligned} R^*R_0^{-1} &= (2cn_0)^{-1}\{a\ E(\hat{\theta}_N - \theta)^2 + c\ EN\} \\ &= (2c)^{-1}\{an_0^{-1}\ E(\hat{\theta}_N - \theta)^2 + c\ EN/n_0\}. \end{aligned}$$

Thus the theorem will be proved once we establish

(2.3.9) $$\lim a(cn_0)^{-1}E(\hat{\theta}_N - \theta)^2 = 1$$

For this, clearly it suffices to show that

$$\text{(2.3.10)} \qquad \lim a(cn_0)^{-1} \, E\{(\hat{\theta}_N - \theta)^2[D]\} = 1$$

and

$$\text{(2.3.11)} \qquad \lim a(cn_0)^{-1} E\{(\hat{\theta}_N - \theta)^2[\overline{D}]\} = 0.$$

First consider (2.3.11). Note that by the maximal inequality for reverse martingales, (1.1.8) and (1.1.10),

$$\sup_{n_1 \le n \le n_2} \|n(\hat{\theta}_n - \theta)^2\|_s = O(1), \; s>0.$$

Therefore by the Holder inequality, Lemma 1.1, Lemma 2.3.1 and similar arguments as in the proof of Theorem 1.2 we get, for $s>2$ and $0<h<s-2$,

$$\begin{aligned}
E\{(\hat{\theta}_N - \theta)^2[N \le n_2]\} &= \sum_{n=n_1}^{n_2} E\{(\hat{\theta}_n - \theta)^2[N=n]\} \\
&\le \sum_{n=n_1}^{n_2} \{\|(\hat{\theta}_n - \theta)^2\|_s \, P^{1-1/s}[N=n]\} \\
&\le \{\sum_{n=n_1}^{n_2} E|\hat{\theta}_n - \theta|^{2s}\}^{1/s}\{P^{1-1/s}[N \le n_2]\} \\
&\le \{\sup_{n_1 \le n \le n_2} \|n(\hat{\theta}_n - \theta)^2\|_s\}(\sum_{n=n_1}^{\infty} n^{-s})^{1/s}\{P^{1-1/s}[N \le n_2]\} \\
&= O(c^{(s-1)/\{2s(1+h)\}}) \, O(c^{(s-1)^2/\{2s(1+h)\}}) \\
&= o(c^{1/2}), \text{ since } 0<h<s-2.
\end{aligned}$$

Similar arguments yield

$$E\{(\hat{\theta}_N - \theta)^2[N>n_3]\} = o(c^{1/2}).$$

Thus by last two rates, (2.3.11) is immediate. Note that for (2.3.10), it suffices to show that, for some c_0,

$$\text{(2.3.12)} \qquad \{a(cn_0)^{-1}\{(\hat{\theta}_N - \theta)^2[D]\}: 0<c<c_0\} \text{ is U.I.}$$

and

$$\text{(2.3.13)} \qquad a(cn_0)^{-1}(\hat{\theta}_N - \theta)^2[D] \xrightarrow{D} \chi_1^2$$

where χ_1^2 denotes the chi square distribution with one degree of freedom. Recall that $a(cn_0)^{-1} = b\sigma^{-1} = n_0\sigma^{-2}$. For (2.3.12), it suffices to show that for some s>1,

$$\sup_{o<c<c_0} E\{n_0\sigma^{-2}(\hat{\theta}_N - \theta)^2[D]\}^s<\infty.$$

Note that

$$E\{n_0\sigma^{-2}(\hat{\theta}_N - \theta)^2[D]\}^s \leq (n_0\sigma^{-2})^s\{E \max_{n_2<n\leq n_3}(\hat{\theta}_n - \theta)^{2s}\}$$

$$\leq \text{const.}(n_0\sigma^{-2})^s\{(E \max_{n_2<n\leq n_3} |\bar{J}_n|^{2s}) + E \max_{n_2<n\leq n_3} |r_n^*|^{2s}\} \tag{2.3.14}$$

where $\bar{J}_n = n^{-1}\Sigma J(Z_i,\delta_i)$, and r_n^* is as in Theorem 1.1. Since $\bar{J}_n$ is a reverse martingale, by the maximal inequality,

$$E\{ \max_{n_2<n\leq n_3} |\bar{J}_n|^{2s} = O(n_2^{-s}). \tag{2.3.15}$$

Now using Lemma 1.1, and the usual integral approximation for sums yield

$$\begin{aligned} E \max_{n_2<n\leq n_3} |r_n^*|^{2s} &\leq \sum_{n=n_2}^{n_3} E|r_n^*|^{2s} \\ &\leq \text{const.}\sum_{n=n_2}^{\infty} n^{-2s} \\ &= O(n_2^{1-2s}) \\ &= O(n_0^{1-2s}). \end{aligned} \tag{2.3.16}$$

Now consider the first term in the R.H.S. of (2.3.14). By (2.3.15),

$$(n_0\sigma^{-2})^s\, E(\max_{n_2<n\leq n_3} |\bar{J}_n|^{2s}) = O(1). \tag{2.3.17}$$

Similarly for the second term in the R.H.S. of (2.3.14), using (2.3.16), we obtain

(2.3.18) $$(n_0\sigma^{-2})^s E \max_{n_2<n\le n_3} |r_n^*|^{2s}\}= O(n^{1-s})$$

Since s can be taken greater than one, (2.3.17) and (2.3.18) imply (2.3.12). As for (2.3.13), it follows from the fact that $[n_2<N\le n_3]\longrightarrow 1$, a.s., Anscombe's and Slutsky's Theorems, Corollary 1.1 and Theorem 1.2. □

Proof of the Theorem 2.3.3: By definition of N_c,

$$b(\hat{\sigma}_N+ N^{-h})\le N<b(\hat{\sigma}_{N-1}+(N-1)^{-h})+1.$$

By dividing all sides by n_0, adding (−1) to all sides and multiplying each side by $n_0^{1/2}$, we get

$$n_0^{1/2}(\hat{\sigma}_N/\sigma -1)+\sigma^{-1}n_0^{1/2}N^{-h}\le n_0^{1/2}(N/n_0 -1)$$
$$< n_0^{1/2}(\sigma^{-1}\hat{\sigma}_{N-1}-1) +(N-1)^{-h}n_0^{1/2}/\sigma +n_0^{-1/2}.$$

By (3.2.1) and since $h>1/2$, the limiting distribution of $n_0^{1/2}(N/n_0- 1)$ is the same as that of $n_0^{1/2}(\hat{\sigma}_N/\sigma - 1)$. From (1.1.11) we obtain

$$n_0^{1/2}(\hat{\sigma}_N - \sigma)\xrightarrow{D} N(0,\gamma^*/(4\sigma^2))$$

which is equivalent to

$$n_0^{1/2}(\hat{\sigma}_N/\sigma - 1)\xrightarrow{D} N(0,\gamma^*/(4\sigma^4))$$

which implies (2.3.4). By taking the square root transformation, we obtain

$$N^{1/2}- n_0^{1/2} \xrightarrow{D} N(0,\gamma^*/(16\sigma^4))$$

which completes the proof of Theorem 2.3.3. □

Chapter 3

Sequential fixed width confidence interval for functionals of the survival curve under random censorship

3.1 Model

Suppose a random sample of size n, $\{(Z_i,\delta_i): 1\leq i\leq n\}$ has been observed. We wish to construct a confidence interval I_n for $\theta = \theta(F) = \int\psi dF$ and $\theta = \int F d\psi$, of prescribed width 2d such that, asymptotically as n tends to infinity, the coverage probability is at least $(1-2\alpha)$. We assume $F, G \in \mathbf{F}$ and $\psi\in\Psi_1$. Note that by Remark 2.3.1, it suffices to consider $\theta = \int\psi dF$ and ψ as in Assumption 1.1.

Notice that for each n an appropriate estimator of θ is $\hat{\theta}_n = \int\psi d\hat{F}_n$, where $\hat{F}_n$ is the P–L estimator of F.

In the rest of this chapter all unspecified limits are considered as d tends to zero. For a given positive real number d and $\alpha\in(0,1/2)$, in view of (1.1.9), let us take $I_n = (\hat{\theta}_n-d,\ \hat{\theta}_n+d)$ with $n = n_d$ defined by

$$n_d = \inf\{k\geq 1: k \geq d^{-2} z_\alpha^2 \sigma^2\} \tag{3.1.1}$$

where z_γ is the upper 100γ percentage point of the standard normal distribution. Then we have

$$\lim P\,[\theta\in I_{n_d}] = 1-2\alpha$$

and

$$\lim \{n_d d^2 z_\alpha^{-2}\sigma^{-2}\} = 1.$$

Since F and G are unknown, the specification of the "optimal" sample size in (3.1.1) cannot be made. We are therefore led to construct a sequential procedure in which the sample size is a positive integer valued random variable $N = N_d$, and the desired confidence interval for θ is

$I_N = (\hat{\theta}_N - d, \hat{\theta}_N + d)$. Motivated by (3.1.1), we define the stopping time $N = N_d$, by

(3.1.2) $$N_d = \inf\{k \geq n_{1d}: k \geq b(\hat{\sigma}_k^2 + k^{-h})\}$$

where $b = d^{-2} z_\alpha^2$ and h is a positive constant. Since $N_d \geq b\, N_d^{-h}$, a.s., we may assume $n_1 = n_{1d} = b^{1/(1+h)}$. Note that n_d, the optimal sample size, is asymptotically equivalent to $n_0 = n_{0d} = b\sigma^2$, that is, $n_d n_{0d}^{-1} \to 1$.

In the rest of this chapter we shall drop the subscript d in N_d, n_d, n_{0d} and on various entities when there is no possibility of confusion. All unspecified limits are considered as d tends to zero or b tends to infinity.

3.2 Main results of this chapter:

The following results hold under Assumption 1.1.

Theorem 3.2.1: For each positive real number h and $F, G \in \mathbf{F}$, (N, I_N) is both consistent and efficient. In fact we shall show that for each $P \in \mathbf{P}$,

(3.2.1) $$\lim P[\theta \in I_N] = 1 - 2\alpha$$

and

(3.2.2) $$\lim E|Nn_0^{-1} - 1| = 0.$$

Theorem 2.2: Let $h > 1/2$ and $F, G \in \mathbf{F}$, then

(3.2.3) $$n_0^{1/2}(n_0^{-1}N - 1) \xrightarrow{D} N(0, \gamma^*/\sigma^4)$$

or equivalently

(3.2.4) $$N^{1/2} - n_0^{1/2} \xrightarrow{D} N(0, \gamma^*/(4\sigma^4))$$

where γ^* is as in (1.1.12).

Proof of Theorem 3.2.1: By definition of N_d, $\lim N_d = \infty$, a.s.. Furthermore if $0 < d_1 < d_2$, we have $N_{d_1} \geq N_{d_2}$ a.s., that is, N_d is nondecreasing as d decreases.

Note that it follows from the representation (1.1.8) that $\{\hat{\theta}_n\}$ is strongly consistent estimator of θ and we have seen in Chapter 2 that $\{\hat{\sigma}_n^2\}$ is a strongly consistent estimator of σ^2. Since $N\uparrow\infty$, a.s.,

$$\hat{\theta}_N \to \theta \quad \text{a.s.,}$$

and

$$\hat{\sigma}_N^2 \to \sigma^2 \quad \text{a.s..} \tag{3.2.5}$$

From (3.1.1) and (3.1.2), and for d sufficiently small,

$$z_\alpha^2\, \sigma^2 \le d^2 n_d < d^2 + z_\alpha^2\, \sigma^2 \tag{3.2.6}$$

and

$$b\, n_0^{-1}\, \hat{\sigma}_N^2 \le N/n_0 \le b\, n_0^{-1}\, \{\, \hat{\sigma}_{N-1}^2 + (N-1)^{-h}\, \} \tag{3.2.7}$$

whence (3.2.6), (3.2.7) together with (3.2.5) yield

$$d^2 n_0 \to z_\alpha^2\, \sigma^2$$

and

$$n_0^{-1} N \to 1 \quad \text{a.s..} \tag{3.2.8}$$

To show (3.2.2), we need a rate on the tail behavior of the stopping time N_d. Since the method of obtaining this rate is analogous to Lemma 2.3.1, we state a similar lemma without proof.

Lemma 3.2.1: For each $0<\epsilon<1$, and any $r<\infty$,

$$P\,[N \le n_0(1-\epsilon)] = O(d^{2(r-1)/(1+h)}) \tag{3.2.9}$$

and

$$\sum_{n \ge n_3} P[N>n] = O(d^{2(r-1)}) \tag{3.2.10}$$

where $n_2 = n_{2d} = \text{int}(n_0(1-\epsilon))$ and $n_3 = n_{3d} = \text{int}(n_0(1+\epsilon))$.

Now by (3.2.8) and Lemma 3.2.1 and arguments similar to those used in the proof of the Theorem 2.3.1, (3.2.2) obtains. From representation (1.1.8), Theorem 1.2, Anscombe's Theorem and (1.1.9), it follows that,

$$N^{1/2}(\hat{\theta}_N - \theta) \xrightarrow{D} N(0,\sigma^2) \tag{3.2.11}$$

from which we get

$$P[\theta \in I_N] = P[\hat{\theta}_N - d \le \theta \le \hat{\theta}_N + d]$$

$$= P[N^{1/2}|\hat{\theta}_N - \theta| \le d\, N^{1/2}]. \tag{3.2.12}$$

Since $d\, N^{1/2} \longrightarrow \sigma z_\alpha$, a.s., (3.2.12), (3.2.11) and Slutsky's theorem establish (3.2.1). □

Proof of Theorem 3.2.2: From (3.1.1), (3.1.2) and similar arguments as the one used in the proof of the Theorem 2.3.3, show that the asymptotic distribution of $n_0^{1/2}(Nn_0^{-1} - 1)$ is the same as that of $n_0^{1/2}(\hat{\sigma}_N^2 - \sigma^2)/\sigma^2$, from which (3.2.3) is immediate. Now (3.2.4) follows from (2.3.3) by square root transformation, that is,

$$n_0^{1/2}(N^{1/2}n_0^{-1/2} - 1) \xrightarrow{D} N(0,\gamma^*/(4\sigma^4)),$$

and the converse is similar. This completes the proof of Theorem 3.2.2. □

Chapter 4

Sequential fixed width confidence interval for a location parameter under random censorship using M–estimation

4.1 Model

Let ϵ and Y be independent random variables and

$$X = \Delta + \epsilon, \quad \Delta \epsilon \Theta \tag{4.1.1}$$

where Θ is an open subset of the real line **R** with compact closure.

The following notations will be used only in this chapter. Let $F(t) = P[\epsilon>t]$, $G(t) = P[Y>t]$, $F_\Delta = F(\cdot-\Delta)$, $H_\Delta = F_\Delta G$, $\psi_\Delta = \psi(\cdot-\Delta)$, $\tilde{\tilde{H}}_\Delta(t) = -\int_{-\infty}^{t} G dF_\Delta$, $\{X_i: i\geq 1\}$ be iid rv's with the same distribution function as X, $\{Y_i: i\geq 1\}$ be iid rv's, independent of $\{X_i: i\geq 1\}$, with the same distribution function as Y, and (1–F), (1–G) are continuous distribution functions and all of the unspecified integrals are on the whole real line.

When dealing with survival time data, one can take X_i's to be $\log_{10}$ or ln of the survival times. The problem considered in this chapter is the sequential interval estimation of Δ using M–estimation based on $\{(Z_i,\delta_i): i\geq 1\}$ where $Z_i = \min(X_i,Y_i)$ and $\delta_i = [X_i \leq Y_i]$.

Let $\hat{F}_n$ be the P–L estimator based on $\{(Z_i,\delta_i): 1\leq i\leq n\}$. An M–estimator of Δ is defined as the solution in t of

$$\lambda_n(t) = \int \psi(x-t) d\hat{F}_n(x) = 0 \tag{4.1.2}$$

for some given function ψ. In the absence of censorship, if $\psi(x,\Delta) = -\frac{\partial}{\partial t}\log f(x-t)|_{t=\Delta}$, where f is the density of the measure induced by X on **R** with respect to the Lebesgue measure, then the solution to (4.1.2), is the Maximum Likelihood Estimate (MLE). Huber (1964) proposed M–estimation as a generalization of (MLE), with desirable robustness properties. Two important examples which are mostly used for the problem of locating the center of a symmetric distribution, say Δ, are the Huber M–estimate by taking Huber ψ function defined by

(4.1.3) $$\psi(x) = \{(-T \vee x) \wedge T\}$$

where T is a positive constant and Tukey's biweight

(4.1.4) $$\psi(x) = x(1-x^2)^2 \, [|x| \leq 1]$$

in which case the defining equation becomes

$$\int \psi(x-t) dF_{\Delta}(x) = 0.$$

In practice it is usually necessary to estimate the scale parameter of the underlying distribution, but this will not be considered in this thesis.

4.2 Assumptions and some preliminary results:

Assumption (A1): F is symmetric about zero and F,G are continuous.

Assumption (A2): Let M and T be constants such that $|\Delta| \leq M$, for all $\Delta \epsilon \Theta$, $G(T+M)>0$ and $F(T)>0$. Let ψ be a monotone nondecreasing, continuous, skew symmetric function and has two continuous bounded derivatives ψ', ψ'' on $(-T,T)$ and ψ is constant on $\{x: x \geq T\} \cup \{x: x \leq -T\}$.

Assumption (A3): $\gamma = \int \psi' dF \neq 0$.

Assumption (A4): $t = \Delta$, is an isolated root of the equation

(4.2.1) $$\lambda_F(t) = \int \psi(x-t) dF_{\Delta}(x) = 0.$$

Remark 4.2.1: For nondecreasing ψ, $\hat{\Delta}_n$ may be written as

$$\hat{\Delta}_n = 1/2 \, (\sup\{t: \lambda_n(t)<0\} + \inf\{t: \lambda_n(t)>0\}).$$

The next lemma is similar to Lemma 7.2.A of Serfling (1980) which has been considered for the case of no censoring. Now we are following the same lines of proof to get similar results in the presence of censorship.

Lemma 4.2.1: Under Assumptions (A2) and (A4), a sequence of solutions $\{\hat{\Delta}_n\}$ to the equation (4.1.2) exists and converges to Δ, a.s..

Proof: Let ϵ be a given positive real number and $\hat{\Delta}_n$ as in Remark 4.2.1. Then $\lambda_F(\Delta-\epsilon)<0<\lambda_F(\Delta+\epsilon)$. We will show that $\lambda_n(t) \rightarrow \lambda_F(t)$, a.s., for each

t for which $|t|\le M$. Assumption (A2) and integration by parts yields

$$|\int\psi(x-t)d\hat{F}_n(x) - \int\psi(x-t)dF_\Delta(x)| = |\int(\hat{F}_n(x) - F_\Delta(x))\, d\psi(x-t)|.$$

Note that, for any $c<\tau_{H_\Delta}$,

$$\sup_{t\le c}|\hat{F}_n(t) - F(t)| \longrightarrow 0, \text{ a.s..}$$

Thus for all t such that $|t|\le M$, we conclude that $\lambda_n(t) \longrightarrow \lambda_F(t)$, a.s..

Therefore

$$P[\Delta-\epsilon\le\hat{\Delta}_n\le\Delta+\epsilon;\text{ i.o.}] = P[\lambda_n(\Delta-\epsilon)\le 0 \text{ and } \lambda_n(\Delta+\epsilon)\ge 0;\text{ i.o.}] = 1$$

where i.o. stands for infinitly often. Since ϵ is arbitrary,

$$\hat{\Delta}_n \longrightarrow \Delta \quad \text{a.s..} \quad \square$$

Lemma 4.2.2: Under Assumption (A2) and for $\{T_n\}$ a sequence of random variables such that $|T_n|\le M$, a.s., and $T_n \longrightarrow \Delta$, a.s., where M is as in Assumption (A2),

$$\int\psi(x-T_n)d\hat{F}_n(x) \longrightarrow \int\psi(x-\Delta)dF_\Delta(x), \text{ a.s..} \tag{4.2.2}$$

Proof: Using the triangle inequatity, Taylor's expansion and integration by parts we have

$$\begin{aligned}
&|\int\psi(x-T_n)d\hat{F}_n(x) - \int\psi(x-\Delta)dF_\Delta(x)| \\
&\le|\int\psi(x-T_n)d(\hat{F}_n(x) - F_\Delta(x))|+|\int\{\psi(x-\Delta) - \psi(x-T_n)\}dF_\Delta(x)| \\
&= |\int(\hat{F}_n(x) - F_\Delta(x))d\psi(x-T_n)| + |\int(T_n-\Delta)\psi'(x-\tilde{\tilde{T}}_n)dF_\Delta(x)|
\end{aligned}$$

where $\tilde{\tilde{T}}_n$ is between T_n and Δ and $\psi'(x-\tilde{\tilde{T}}_n) = \frac{\partial}{\partial t}\psi(x-t)|_{t=\tilde{\tilde{T}}_n}$. Since ψ, ψ' are bounded, $T_n \longrightarrow \Delta$, a.s., and

$$\sup_{x\le T+M}|\hat{F}_n(x) - F_\Delta(x)| \longrightarrow 0 \quad \text{a.s.,}$$

the lemma is immediate. $\square$

Remark 4.2.3 The results of our previous chapters which are given for life times (nonnegative random variables) with or without the restriction of $T<\tau$, can be extended to the case of any random variable with similar restrictions on the distributions (1–F), (1–G) and the function ψ.

Theorem 4.2.1: Suppose Assumptions (A1) through (A4) hold and let $\hat{\Delta}_n$ be a solution sequence of (4.1.2), then

$$n^{1/2}(\hat{\Delta}_n - \Delta) \xrightarrow{D} N(0,\sigma_\Delta^2)$$

where

$$\sigma_\Delta^2 = \sigma_\Delta^2(F,G) = \gamma^{-2}\int A_\Delta^2 \, H_\Delta^{-2} \, d\tilde{\tilde{H}}_\Delta \tag{4.2.4}$$

and

$$A_\Delta(t) = \int_t^\infty F_\Delta d\psi_\Delta, \quad t\in\mathbb{R}.$$

Proof: Since ψ is differentiable, so is the function

$$\lambda_n(t) = \int\psi(x-t)d\hat{F}_n = \Sigma\psi(Z_i-t)\delta_i d_{in}$$

where d_{in} is the jump of the P–L estimator at Z_i. Therefore we have

$$\int\psi(x-\hat{\Delta}_n)d\hat{F}_n(x) - \int\psi(x-\Delta)d\hat{F}_n(x) = \int(\hat{\Delta}_n - \Delta)\psi'(x-\tilde{\tilde{\Delta}}_n)d\hat{F}_n(x)$$

where $|\tilde{\tilde{\Delta}}_n - \Delta|\le|\hat{\Delta}_n - \Delta|$ and $\psi'(x-\tilde{\tilde{\Delta}}_n) = \frac{\partial}{\partial t}\psi(x-t)|_{t=\tilde{\tilde{\Delta}}_n}$. Since $\lambda_n(\hat{\Delta}_n) = \lambda_F(\Delta) = 0$, integration by parts and some algebra yield

$$\begin{aligned} n^{1/2}(\hat{\Delta}_n - \Delta) &= \tilde{\tilde{\gamma}}_n^{-1}\{n^{1/2}\int(\hat{F}_n - F_\Delta)d\psi_\Delta\}[\tilde{\tilde{\gamma}}_n\neq 0] \\ &= \tilde{\tilde{\gamma}}_n^{-1}n^{1/2}L_n[\tilde{\tilde{\gamma}}_n\neq 0] \end{aligned}$$

where $\tilde{\tilde{\gamma}}_n = \int\psi'(x-\tilde{\tilde{\Delta}}_n)d\hat{F}_n(x)$ and $L_n = \int(\hat{F}_n - F_\Delta)d\psi_\Delta$.

Recall the representation (1.1.8) and note that under Assumptions (A2) and (A3) and similar arguments as is used in the proof of Lemma 4.2.2 yield $\tilde{\tilde{\gamma}}_n \to \gamma$, a.s., and $[\tilde{\tilde{\gamma}}_n\neq 0]\to 1$, a.s.. Thus by Slutsky's Theorem (4.2.3) is achieved. □

Remark 4.2.2: Under the Assumption (A1), Δ is the median of (1–F). Gardiner, Susarla and van Ryzin (1985a) estimated Δ by the sample median,

$\hat{m} = \hat{F}_n^{-1}(1/2) = \inf\{t: \hat{F}_n(t) \le 1/2\}$ which is shown to have the following properties:

(i) For each positive constant s,

(4.2.5) $$\|\hat{m} - \Delta\|_s = O(n^{-1/2})$$

(ii) If F_Δ has a density f_Δ and is positive at Δ, then

(4.2.6) $$n^{1/2}(\hat{m} - \Delta) \xrightarrow{D} N(0, C_\Delta(\Delta)/(4\, f_\Delta^2(\Delta)))$$

where $C_\Delta(t) = \int_{-\infty}^{t} H_\Delta^{-2} d\overset{\approx}{H}_\Delta$. Also they have suggested the use of $C_n(\hat{m})/(4\hat{f}_n^2(\hat{m}))$ as an estimator of $C_\Delta(\Delta)/(4f_\Delta^2(\Delta))$, where $C_n(t) = \int_{(-\infty,t]} (H_n + 1/n)^{-2} d\overset{\approx}{H}_n$, and $\hat{f}_n$ is a suitable estimator of the density f_Δ.

In this chapter our results do not require the existence of a density. However we use $\hat{m} = \hat{F}_n^{-1}(1/2)$, as a preliminary estimate of Δ in the estimation of the asymptotic variance σ_Δ^2. Let

(4.2.7) $$\sigma_n^2(\hat{m}) = (\int \psi'(x-\hat{m}) d\hat{F}_n(x))^{-2} \int \hat{A}_n^2 dC_n$$

where

$$\hat{A}_n(t) = \int_t^\infty \hat{F}_n(s) d\psi(s-\hat{m}).$$

Remark 4.2.3 In the absence of the censoring, Carroll (1978) has an almost sure expansion for M–estimates. In his paper he states that for (4.2.7) to be a consistent estimate of σ_Δ^2, we need F to be symmetric and ψ skew symmetric. Reid (1981) calculated the influence curve for Kaplan–Meier M–estimate and has found the above asymptotic normality by the influence function approach.

4.3 Sequential fixed width confidence interval for the location parameter Δ

In the rest of this chapter all the limits are considered as d tends to zero. In view of (4.2.3), for a given positive real number d and $\alpha \in (0,1/2)$ we take $I_n = (\hat{\Delta}_n - d, \hat{\Delta}_n + d)$ with $n = n_d$ defined by

$$(4.3.1) \qquad n_d = \inf\{k\geq 1:\ k \geq b\ \sigma_\Delta^2\}$$

where $b = d^{-2}z_\alpha^2$. Then we have

$$\lim P[\Delta\in I_n] = 1-2\alpha$$

and

$$\lim \{nb^{-1}\sigma_\Delta^{-2}\} = 1.$$

Since F, G and Δ are unknown, the specification of the "optimal" sample size in (4.3.1) cannot be made. We are therefore led naturally to construct a sequential procedure in which the sample size is a stopping time $N = N_d$

$$(4.3.2) \qquad N_d = \inf\{k\geq n_1:\ k \geq b(\hat{\sigma}^2(\hat{m})+k^{-h})\}$$

where n_1 is as in (3.1.2) and $\hat{\sigma}^2(\hat{m})$ is given by (4.2.7).

Before we present the properties of the above sequential procedure and the stopping time, we provide some preliminary results.

The following lemma is stated in Sriram (1987).

Lemma 4.3.1 (Lemma 1 of Sriram (1987)): Let U_n, V_n be any sequence of random variables and a, b $\neq$ 0 and s>0 be real numbers. If

$$P[|U_n-a|\geq\epsilon] = O(n^{-s}) = P[|V_n-b|\geq\epsilon], \text{ for every } \epsilon>0,$$

then

$$P[|U_n/V_n - a/b|\geq\epsilon] = O(n^{-s}), \text{ for every } \epsilon>0.$$

Now we shall show that for each positive ϵ and some r>1,

$$(4.3.3) \qquad P[|\hat{\sigma}_n^2(\hat{m}) - \sigma_\Delta^2| \geq \epsilon] = O(n^{-r}).$$

Note that by Lemma 4.3.1 and the Assumption (A3), it suffices to show that

$$(4.3.4) \qquad P[|\int\psi'(x-\hat{m})d\hat{F}_n(x) - \gamma| \geq \epsilon] = O(n^{-r})$$

and

$$(4.3.5) \qquad P[|\int\hat{A}_n^2 dC_n - \int A_\Delta^2 H_\Delta^{-2} d\tilde{\tilde{H}}_\Delta| \geq \epsilon\ ] = O(n^{-r})$$

First consider (4.3.4). By Taylor's expansion

$$\int\psi'(x-\hat{m})d\hat{F}_n(x) = \int\psi'(x-\Delta)d\hat{F}_n(x) + (\hat{m}-\Delta)\int\psi''(x-\tilde{\Delta}_n)d\hat{F}_n(x)$$

where $\tilde{\Delta}_n$ is between Δ and $\hat{m}$. Now by the above expansion, Assumption (A2) and integration by parts we have

$$E|\int\psi'(x-\hat{m})d\hat{F}_n(x) - \gamma|^{2r} \le \text{const.}\{E|\hat{m}-\Delta|^{2r} + \int E|\hat{F}_n - F_\Delta|^{2r} d\psi'(x-\Delta)\}.$$

Now by (4.2.5), the representation (1.1.8), Lemma 1.1 and Marcinkiewicz–Zygmund inequality we have

$$E|\int\psi'(x-\hat{m})d\hat{F}_n(x) - \gamma|^{2r} = O(n^{-r})$$

which, by Markov's inequality, implies (4.3.4). A similar calculation to that done in the proof of Theorem 1.4, which is shown in Appendix A, and consideration of the extension explained in Remark 4.2.3, leads to (4.3.5). This completes the proof of (4.3.3).

Recall that n_d, the "optimal" sample size, is asymptotically equivalent to $n_0 = n_{0d} = b\sigma_\Delta^2$. In the following, whenever there is no possibility of confusion, the subscript d of N_d, n_d and n_{0d} will be dropped. Now we state the main results of this chapter.

Theorem 4.3.1 Under Assumptions (A1) through (A4) and for $h>0$, (N, I_N) is both consistent and efficient, that is, for each $\Delta\epsilon\Theta$ and $P\epsilon P$,

(4.3.6) $$\lim P[\Delta\epsilon I_N] = 1-2\alpha$$

and

(4.3.7) $$\lim E|N\, n_0^{-1} - 1| = 0$$

Proof: Note that by definition of N_d, $N_d \to \infty$, a.s.. Furthermore it can be shown that N_d is nondecreasing as d decreases (see Chapter 2 or 3 for similar arguments). Therefore by Lemma 4.2.1,

$$\hat{\Delta}_N \to \Delta, \text{ a.s.,}$$

and (4.3.3) together with Borel–Cantelli lemma yield

$$\hat{\sigma}_N^2(\hat{m}) \to \sigma_\Delta^2, \text{ a.s..}$$

From the definition of N_d and n_d and arguments similar to those in Chapter 3 we obtain, for d sufficiently small, (4.3.8) $$n_0^{-1} N \longrightarrow 1 \quad \text{a.s.,}$$

and

(4.3.9) $$d^2 n_0 \longrightarrow z_\alpha^2 \sigma_\Delta^2$$

Note that, as in Chapter 3, for showing (4.3.6) we need to show that

(4.3.10) $$N^{1/2}(\hat{\Delta}_N - \Delta) \xrightarrow{D} N(0,\sigma_\Delta^2).$$

Recall that

$$N^{1/2}(\hat{\Delta}_N - \Delta) = \tilde{\tilde{\gamma}}_N^{-1} N^{1/2} L_N \, [\tilde{\tilde{\gamma}}_N \neq 0]$$

and

$$L_N = \int(\hat{F}_N - F_\Delta)d\psi_\Delta = -N^{-1}\Sigma J_\Delta(Z_i,\delta_i) + r^*_{N,\Delta}.$$

where J_Δ and $r^*_{N,\Delta}$ are the same as J and r^*_n, given by (1.1.4) and Theorem 1.1, respectively when F is replaced by F_Δ. Now write

$$N^{1/2}(\hat{\Delta}_N - \Delta) = \{(\tilde{\tilde{\gamma}}_N^{-1} - \gamma^{-1})\, N^{1/2} r^*_{N,\Delta} + \gamma^{-1} N^{1/2} r^*_{N,\Delta}$$
$$- \gamma^{-1} N^{-1/2}\, \bar{J}_{N,\Delta} - (\tilde{\tilde{\gamma}}_N^{-1} - \gamma^{-1})\, N^{-1/2}\, \bar{J}_{N,\Delta}\}[\tilde{\tilde{\gamma}}_N \neq 0]$$

where

$$\bar{J}_{N,\Delta} = N^{-1}\Sigma J_\Delta(Z_i,\delta_i).$$

Therefore by (1.1.8), Anscombe and Slutsky's Theorems, (4.3.10) is immediate and so is (4.3.6). To show (4.3.7) we need similar rates as the one given in Lemma 3.2.1. Note that all we need to get such rates is (4.3.3). Hence, following the same lines of proof of Lemma 3.2.1 and Theorem 3.2.1, we obtain (4.3.7) which completes the proof of Theorem 4.3.1. □

Remark 4.3.1 The asymptotic normality of the stopping time N_d, can be obtained by arguments similar to those given earlier in Chapters 2 & 3 but obtaining the exact form of the asymptotic variance is an immersely tedious calculation. Thus one can establish

$$n_0^{1/2}(N_d\, n_0^{-1} - 1) \xrightarrow{D} N(0,\beta)$$

and

$$N_d^{1/2} - n_0^{1/2} \xrightarrow{D} N(0,\beta/4)$$

but the exact computation of β is difficult.

APPENDICES

Appendix A

The almost sure representation and the asymptotic distribution of the estimator of the asymptotic variance of the estimator of functionals of the form $\theta(F)=\int\psi dF$ and $\theta(F)=\int F d\psi$, of survival curve F under random censorship

We shall present here the proof of Theorem 1.3. To the best of our knowledge this is a new result. In the sequel $\psi\in\Psi_1$ and all the limits are considered as n tends to infinity. Note that by Remark 2.3.1, it suffices to consider ψ as in Assumption 1.1. Recall that

$$\Gamma_{ij}(t) = \int_0^t A^i dC^j,\ t \geq 0,\ i,j = 1,2,3,4,$$

$$K(t) = 1 + \Sigma\ [Z_i>t],\quad C(t) = \int_0^t H^{-2} d\tilde{\tilde{H}},\ t<\tau,$$

$$A(t) = \int_{[t,\infty)} F d\psi,\ A_n(t) = \int_{[t,\infty)} \hat{F}_n d\psi,\quad 0\leq t<\infty$$

and

$$\hat{\sigma}_n^2 = \int A_n^2 n^2 K^{-2} d\tilde{\tilde{H}}_n,\quad \sigma^2 = \int A^2 H^{-2} d\tilde{\tilde{H}}.$$

Proof of Theorem 1.3:

To simplify notation let $\varphi_n = n^2 A_n{}^2 K^{-2}$ and $\varphi = A^2\ H^{-2}$. We write

$$n^{1/2}(\hat{\sigma}_n^2 - \sigma^2) = n^{1/2}(\int\varphi_n d\tilde{\tilde{H}}_n - \int\varphi d\tilde{\tilde{H}}\ )$$

$$= n^{1/2}\{\int(\varphi_n - \varphi)d(\tilde{\tilde{H}}_n - \tilde{\tilde{H}}) + \int(\varphi_n - \varphi)d\tilde{\tilde{H}} + \int\varphi d(\tilde{\tilde{H}}_n - \tilde{\tilde{H}})\}$$

$$= n^{1/2}(\ D_1 + D_2 + D_3\ ),\ \text{say.} \tag{A.1}$$

Note that

$$\varphi_n - \varphi = (\varphi_n^{1/2} - \varphi^{1/2})^2 + 2\ \varphi^{1/2}(\varphi_n^{1/2} - \varphi^{1/2})$$

and

$$\varphi_n^{1/2} - \varphi^{1/2} = nK^{-1}(A_n - A) + A(nK^{-1} - H^{-1})$$
$$= H^{-1}(A_n - A) + A(nK^{-1} - H^{-1}) + (A_n - A)(nK^{-1} - H^{-1})$$
$$= B_n + R_{n,1}$$

where $B_n = H^{-1}(A_n - A) + A(nK^{-1} - H^{-1})$ and $R_{n,1} = (A_n - A)(nK^{-1} - H^{-1})$.

Therefore

$$\varphi_n - \varphi = (B_n + R_{n,1})^2 + 2AH^{-1}(B_n + R_{n,1}).$$

We simplify to get

$$\text{(A.2)} \qquad \varphi_n - \varphi = 2AH^{-2}\{(A_n - A) - AH^{-1}(H_n - H)\} + R_{n,2}$$

where

$$R_{n,2} = (B_n + R_{n,1})^2 + 2AH^{-1}R_{n,1} + 2A^2H^{-3}(H_n - H) - 2A^2H^{-1}(H^{-1} - nK^{-1}).$$

Therefore it follows from (A.1) and (A.2) that

$$D_2 = 2\int(A_n - A)d\Gamma_{11} - 2\int H^{-1}(H_n - H)d\Gamma_{21} + \int R_{n,2}d\tilde{\tilde{H}}$$
$$= 2\,U_1 - 2\,U_2 + R_{n,3}\ , \text{ (say)}.$$

Note that we can rewrite U_1 and U_2 as

$$U_1 = \int\{\int_{\cdot}^{\infty}(\hat{F}_n - F)d\psi\}d\Gamma_{11}$$

and

$$U_2 = n^{-1}\Sigma\int A^2H^{-3}\{[Z_i > \cdot] - H\}\ d\tilde{\tilde{H}}$$
$$= n^{-1}\Sigma\int H^{-1}\{[Z_i > \cdot] - H\}\ d\Gamma_{21}.$$

Recall the representation (1.1.5). We rewrite U_1 as

$$U_1 = U_3 + R_{n,4}$$

where

$$U_3 = n^{-1}\Sigma\int\{\int_{\cdot}^{\infty}\xi(Z_i,\delta_i;s)F(s)d\psi(s)\}d\Gamma_{11}$$

and

$$R_{n,4} = \int\left(\int_{\cdot}^{\infty} r_n F d\psi\right) d\Gamma_{11}$$

and r_n is as in (1.1.5). Hence

$$n^{1/2}(\hat{\sigma}_n^2 - \sigma^2) = 2n^{1/2}U_3 - 2n^{1/2}U_2$$

(A.3)
$$+ n^{-1/2}\Sigma\{A^2(Z_i)H^{-2}(Z_i)\delta_i - \sigma^2\} + n^{1/2}R_{n,5}$$

where $R_{n,5} = D_1 + R_{n,3} + 2\,R_{n,4}$. After some algebra on (A.3) we get

$$n^{1/2}(\hat{\sigma}_n^2 - \sigma^2) = n^{-1/2}\Sigma V_i + n^{1/2}R_{n,5}$$

where

$$V_i = 2W_{i,1} + W_{i,2},$$

$$W_{i,1} = \int\left\{\int_{\cdot}^{\infty} \xi(Z_i,\delta_i;s)F(s)d\psi(s)\right\}d\Gamma_{11}$$

and

$$W_{i,2} = A^2(Z_i)H^{-2}(Z_i)\delta_i - 2\int H^{-1}[Z_i > \cdot]d\Gamma_{21} + \sigma^2.$$

Under assumptions of Theorem 1.3, $\{V_i: i \geq 1\}$ is a sequence of bounded iid rv's with mean

$$EV_1 = EW_{1,1} = EW_{1,2} = 0.$$

To obtain the almost sure representation of $\hat{\sigma}^2$ we need to show that

$$n^{1/2}R_{n,5} \rightarrow 0 \quad \text{a.s.},$$

in addition if we show that the variance of V_1 is γ^*, then (1.1.11) will be achieved by the central limit theorem. First we compute the variance of V_1. Let (Z,δ), V, W_1 and W_2 be copies of (Z_1,δ_1), V_1, $W_{1,1}$ and $W_{1,2}$ respectively, then

(A.4) $$\text{Var } V = E(2W_1 + W_2)^2 = 4EW_1^2 + EW_2^2 + 4EW_1W_2.$$

In the following we use repeatedly the Fubini Theorem, integrations by parts and the identities

$$(\int g d\tilde{\tilde{H}})^2 = 2\int g(t)(\int_0^t g d\tilde{\tilde{H}})d\tilde{\tilde{H}}(t),$$

$$Eg(Z) = \int g d(-H)$$

and

$$E\{g(Z)\delta\} = \int g d\tilde{\tilde{H}}.$$

Now we are ready to compute EW_2^2. Since W_2 has mean zero,

$$E\{A^2(Z_i)H^{-2}(Z_i)\delta_i - 2\int H^{-1}[Z_i>\cdot]d\Gamma_{21}\} = -\sigma^2.$$

Therefore

$$EW_2^2 = E\{A^4(Z)H^{-4}(Z)\delta\} + 4E(\int H^{-1}[Z>\cdot]d\Gamma_{21})^2$$
$$- 4E\{A^2(Z)H^{-2}(Z)\delta\int H^{-1}[Z>\cdot]d\Gamma_{21}\} - \sigma^4.$$

Note that

$$E(\int H^{-1}[Z>\cdot]d\Gamma_{21})^2 = 2E\{(\int(\int_0^s H^{-1}[Z>\cdot]d\Gamma_{21})H^{-1}(s)[Z>s]d\Gamma_{21}(s)\}$$
$$= 2\int(\int_0^{\cdot} H^{-1}d\Gamma_{21})d\Gamma_{21}$$
$$= 2\int H^{-1}(\int_{\cdot}^{\infty} d\Gamma_{21})d\Gamma_{21}$$

and

$$E\{A^2(Z)H^{-2}(Z)\delta\int H^{-1}[Z>\cdot]d\Gamma_{21}\} = \int H^{-1}(\int_{\cdot}^{\infty} d\Gamma_{21})d\Gamma_{21}.$$

We simplify to obtain

$$\text{(A.5)} \qquad EW_2^2 = \int H^{-2}d\Gamma_{41} + 4\int H^{-1}(\int_{\cdot}^{\infty} d\Gamma_{21})d\Gamma_{21} - \sigma^4.$$

To simplify notation, throughout $\xi(Z,\delta;t)$ will be abbreviated to $\xi(t)$. Since $E\xi(s)\xi(v) = C(s\Lambda v)$ for $s,v<\tau$,

$$EW_1^2 = E(\int\int_{\cdot}^{\infty} \xi(s)F(s)\, d\psi(s)d\Gamma_{11})^2$$
$$= 2E\{\int\{(\int_t^{\infty} \xi F d\psi)\int_0^t (\int_u^{\infty} \xi F d\psi)d\Gamma_{11}(u)\}d\Gamma_{11}(t)\}$$
$$= 2\int\{\int_t^{\infty} F(s)\{\int_0^t \{\int_u^{\infty} C(s\Lambda v)F(v)d\psi(v)\}d\Gamma_{11}(u)\}d\psi(s)\}d\Gamma_{11}(t).$$

Now consider the most inner integral on two sets $[s\le v]$, $[s>v]$ and note that on the set $[s\le v]$, $u\le t\le s\le v$. Hence

$$EW_1^2 = 2\int\{\int_t^\infty ACFd\psi\}\Gamma_{11}(t)d\Gamma_{11}(t)$$
$$+ 2\int\{\int_t^\infty F(s)\{\int_0^t\{\int_u^s CFd\psi\}d\Gamma_{11}(u)\}d\psi(s)\}d\Gamma_{11}(t).$$

Integration by parts on the inner integral of the second term on the R.H.S. of the last equation yields

$$EW_1^2 = \int\Gamma_{22}d\Gamma_{21} + 2\int(\int_\cdot^\infty \Gamma_{11}Fd\psi)\Gamma_{11}d\Gamma_{11} - \int\Gamma_{11}^2 d\Gamma_{21}.$$

Now integration by parts on the middle term of the R.H.S. of the last equation gives

(A.6) $$EW_1^2 = \int(2\Gamma_{11}^2 + \Gamma_{22})d\Gamma_{21}.$$

Now we consider EW_1W_2.

$$EW_1W_2 = E\{A^2(Z)H^{-2}(Z)\delta\int\{\int_\cdot^\infty C(Z\wedge s)F(s)d\psi(s)\}d\Gamma_{11}\}$$
$$- E\{A^2(Z)H^{-3}(Z)\delta\int\{\int_Z^\infty Fd\psi\}d\Gamma_{11}\}$$
$$- 2E\{\int\int_\cdot^\infty C(Z\wedge s)F(s)d\psi(s)d\Gamma_{11}\int H^{-1}[Z>\cdot]d\Gamma_{21}\}$$
$$+ 2E\{\delta H^{-1}(Z)\{\int\{\int_Z^\infty F(s)d\psi(s)\}d\Gamma_{11}\}\int H^{-1}[Z>\cdot]d\Gamma_{21}\}$$

(A.7) $$= E\{Q_1 - Q_2 - 2Q_3 + 2Q_4\}, \text{ say.}$$

To compute EQ_1, consider Q_1 on the two sets $[Z\le s]$ and $[Z>s]$. Therefore, by Fubini Theorem

$$EQ_1 = \int\{\int_\cdot^\infty E(C(Z)A^2(Z)H^{-2}(Z)\delta[Z\le s])F(s)d\psi(s)\}d\Gamma_{11}$$
$$+ \int\{\int_\cdot^\infty E\{A^2(Z)H^{-2}(Z)\delta[Z>s]\}C(s)F(s)d\psi(s)\}d\Gamma_{11}$$
$$= (1/2)\int(\int_\cdot^\infty \Gamma_{22}Fd\psi)d\Gamma_{11} + \int\{\int_\cdot^\infty C(s)(\int_s^\infty d\Gamma_{21})F(s)d\psi(s)\}d\Gamma_{11}.$$

Similar arguments yield

$$EQ_2 = \int\{\int_{\cdot}^{\infty} E(A^2(Z)H^{-3}(Z)\delta[Z\leq s])F(s)d\psi(s)\}d\Gamma_{11}$$
$$= \int\{\int_{\cdot}^{\infty} (\int_0^s H^{-1}d\Gamma_{21})F(s)d\psi(s)\}d\Gamma_{11}.$$

Similar arguments as is used in the computation of EQ_1 yield

$$EQ_3 = E(\int\int_{\cdot}^{\infty} C(Z)[Z\leq s]F(s)d\psi(s)d\Gamma_{11})(\int H^{-1}[Z>\cdot]d\Gamma_{21})$$
$$+ E\ (\int\int_{\cdot}^{\infty} C(s)[Z>s]F(s)d\psi(s)d\Gamma_{11})(\int H^{-1}[Z>\cdot]d\Gamma_{21})$$
$$= \int\int_{\cdot}^{\infty} \{\int_0^s H^{-1}(u)\{\int_u^s Cd(-H)\}d\Gamma_{21}(u)\}F(s)d\psi(s)d\Gamma_{11}$$
$$+ \int\int_{\cdot}^{\infty} C(s)\{\int_0^{\infty} H^{-1}(u)H(sVu)d\Gamma_{21}(u)\}F(s)d\psi(s)\ d\Gamma_{11}.$$

We consider the last term in the last equation on two sets $[s\leq u]$ and $[s>u]$ and simplify to obtain

$$EQ_3= \int\{\int_{\cdot}^{\infty} \{\int_0^s H^{-1}(u)\{C(u)H(u)-C(s)H(s)+\int_u^s HdC\}d\Gamma_{21}(u)\}F(s)d\psi(s)\}d\Gamma_{11}$$
$$+ \int\int_{\cdot}^{\infty} C(s)H(s)\{\int_0^s H^{-1}d\Gamma_{21}\}F(s)d\psi(s)d\Gamma_{11}$$
$$+ \int\{\int_{\cdot}^{\infty} C(s)\{\int_s^{\infty} d\Gamma_{21}\}F(s)d\psi(s)\}d\Gamma_{11}.$$

Finally, similar but simpler arguments yield

$$EQ_4= E(\int\{\int_{\cdot}^{\infty} \delta[Z\leq s]H^{-1}(Z)F(s)d\psi(s)\}d\Gamma_{11})(\int H^{-1}[Z>\cdot]d\Gamma_{21})$$
$$= \int\int_{\cdot}^{\infty} \{\int_0^s H^{-1}(u)\{\int_u^s H^{-1}d\tilde{\tilde{H}}\}d\Gamma_{21}\}F(s)d\psi(s)\}d\Gamma_{11}.$$

By substituting EQ_1 through EQ_4 in (A.7) we obtain

$$EW_1W_2 = -\int\int_{\cdot}^{\infty} \{\int_0^s Cd\Gamma_{21}\}F(s)d\psi(s)d\Gamma_{11}$$
$$\text{(A.8)} \qquad - \int\int_{\cdot}^{\infty} C(s)\{\int_s^{\infty} d\Gamma_{21}\}F(s)d\psi(s)d\Gamma_{11}$$
$$- \int\int_{\cdot}^{\infty} \{\int_0^s H^{-1}d\Gamma_{21}\}F(s)d\psi(s)d\Gamma_{11}.$$

Now substitution of(A.5), (A.6) and (A.8) in (A.4) and some algebra yield

$$EV = 6\int\Gamma_{11}^2 d\Gamma_{21} - 4\int H^{-1}\Gamma_{11}d\Gamma_{31} + \int H^{-2}d\Gamma_{41} - \sigma^4.$$

We need the following lemma to obtain the almost sure representation of $n^{1/2}(\hat{\sigma}_n^2 - \sigma^2)$.

Lemma A.1: Under Assumption 1.1, for continuous F, G, $T<\tau$ and $\alpha<1/2$,

(i) $$n^{\alpha}\|H_n - H\| \longrightarrow 0 \quad \text{a.s.},$$

(ii) $$n^{\alpha}\|\tilde{\tilde{H}}_n - \tilde{\tilde{H}}\| \longrightarrow 0 \quad \text{a.s.},$$

(iii) $$n^{\alpha}\|nK^{-1} - H^{-1}\|_0^T \longrightarrow 0 \quad \text{a.s.},$$

(iv) $$n^{\alpha}\|\hat{F}_n - F\|_0^T \longrightarrow 0 \quad \text{a.s.},$$

(v) $$n^{\alpha}\|A_n - A\|_0^T \longrightarrow 0 \quad \text{a.s.}$$

where $\|\cdot\|$ denotes the sup norm and $\|\cdot\|_0^T$ means the sup is taken over the closed interval [0,T].

Proof: (i) and (ii) follow from Glivenko–Cantelli Theorem. Let $\epsilon>0$. Since H is monotone nonincreasing and H(T)>0, (iii) also follow from Glivenko–Cantelli Theorem. (iv) follows from Theorem 2 of Shorack and Wellner (1986) page 308. To show (v) note that

$$n^{\alpha}\|A_n - A\|_0^T = \|n^{\alpha}\int_{\cdot}^{\infty}(\hat{F}_n - F)d\psi\|_0^T \le \text{const.}\|n^{\alpha}(\hat{F}_n - F)\|_0^T.$$

Hence (v) is implied by (iv). This completes the proof of the lemma. □

Recall that $R_{n,5} = D_1 + R_{n,3} + 2R_{n,4}$. We shall show that

(a) $$n^{1/2}D_1 \longrightarrow 0, \text{ a.s.},$$

(b) $$n^{1/2}R_{n,3} \longrightarrow 0, \text{ a.s.},$$

(c) $$n^{1/2}R_{n,4} \longrightarrow 0, \text{ a.s.}.$$

Now recall the representation (A.2) of $(\varphi_n - \varphi)$. To show (a) holds, we shall show that each term has such a property. Consider the first term, we want to show that

$$n^{1/2}\int 2AH^{-2}(A_n - A)d(\tilde{\tilde{H}}_n - \tilde{\tilde{H}}) \longrightarrow 0, \text{ a.s.}.$$

This follows from Assumption 1.1, Lemma A.1 (ii) and (v) and similar arguments as are used in the proof of Lemma 2 of Lo and Singh (1986). All other terms can be handled similarly. By similar arguments, it can be shown that (b) holds. To show (c) holds note that $R_{n,4} = \int\{\int^{\infty}_{\cdot} r_n F d\psi\}\, d\Gamma_{11}$, where r_n is as in (1.1.5). Therefore it easily follows, from Lemma 1.1, that

$$n^{1/2}R_{n,4} \to 0, \text{ a.s..}$$

Thus we have the almost sure representation of $\hat{\sigma}^2_n$ and by the CLT (1.1.11) follows. □

Appendix B

Rate of convergence of the estimator of the asymptotic variance, $\sigma^2 = \int A^2 H^{-2} d\tilde{\tilde{H}}$

In this section we shall present the proof of Theorem 1.4. We shall show that for each $\epsilon>0$, and $r<\infty$, (1.1.13) holds. Note that it suffices to prove the theorem for $r>1$.

The following arguments are very similar to that of Gardiner and Susarla (1983) Appendix A, except we are giving a shorter proof using the representation (1.1.5) and Lemma 1.1.

Proof of Theorem 1.4: Let us write

$$\begin{aligned}\hat{\sigma}^2_n - \sigma^2 &= \int nK^{-2}(A^2_n - A^2)d\tilde{\tilde{H}}_n + \int A^2(n^2K^{-2} - H^{-2})d\tilde{\tilde{H}}_n + \int A^2H^{-2}d(\tilde{\tilde{H}}_n - \tilde{\tilde{H}})\\ &= T_{n,1} + T_{n,2} + T_{n,3}, \text{ say.}\end{aligned}$$

First we examine $T_{n,2}$. Since $K = 1 + nH_n$,

$$\begin{aligned}H^2 - n^{-2}K^2 &= H^2 - (H_n + n^{-1})^2\\ &= -n^{-2} - 2n^{-1}H - (H_n - H)^2 - 2n^{-1}(H_n - H) - 2H(H_n - H)\end{aligned}$$

so that on substitution we have

$$\text{(B.1)} \qquad |T_{n,2}| \le \sum_{j=1}^{5} |T_{n,2j}|.$$

To handle the terms in (B.1) we shall use the following result for Binomial moments.

Lemma B.1: Let U be a Binomial random variable with parameters (n,p). Then for any $k \geq 1$

$$E(1 + U)^{-k} \leq k!\ (np)^{-k}.$$

Proof: See Koul, Susarla and van Ryzin (1981), Moment Lemma, Page 1283. □

Recall that

$$T_{n,2} = \int A^2(n^2K^{-2}H^{-2})\{-n^{-2}-2n^{-1}H-(H_n-H)^2-2n^{-1}(H_n-H)-2H(H_n-H)\}d\tilde{\tilde{H}}.$$

We want to compute the rth moment of $T_{n,2}$, for $r>1$. Consider the $T_{n,21}$ term. Note that

$$T_{n,21} = -n^{-1}\Sigma\delta_i A^2(Z_i)H^{-2}(Z_i)K^{-2}(Z_i).$$

Therefore

$$E|T_{n,21}|^r \leq n^{-1}\Sigma\ E\{\delta_i A^{2r}(Z_i)H^{-2r}(Z_i)K^{-2r}(Z_i)\}.$$

In the following E_1 stands for a conditional expectation given (Z,δ), and all c_i's are constants may depend only on r. Note that given (Z_j,δ_j), $(K(Z_j)-1)$ is the sum of $(n-1)$ Bernoulli random variables with probability of success of $H(Z_j)$. Thus it follows, from Lemma B.1, that

$$\begin{aligned} E|T_{n,21}|^r &\leq c_1n^{-1}\{E\{\delta A^{2r}(Z)H^{-2r}(Z)E_1(K^{-2r}(Z))\}\} \\ &\leq c_2E\{\delta A^{2r}(Z)H^{-2r}(Z)\{nH(Z)\}^{-2r}\} \\ &= c_2n^{-2r}\int A^{2r}H^{-4r}d\tilde{\tilde{H}}. \end{aligned}$$

Under Assumption 1.1 and for $F, G \in \mathcal{F}$, it follows that the last integral is finite. Hence

$$E|T_{n,21}|^r = O(n^{-2r}). \tag{B.2}$$

Exactly in similar manner we can bound $E|T_{n,22}|$ and show that

$$E|T_{n,22}|^r = O(n^{-r}). \tag{B.3}$$

Recall that

$$T_{n,23} = -n^{-1}\Sigma(n^2K^{-2}A^2H^{-2}(H_n - H)^2)(Z_i)\cdot\delta_i.$$

Thus

$$E|T_{n,23}|^r \le c_3 E\{(n^2K^{-2}A^2H^{-2}(H_n - H)^2)^r(Z)\cdot\delta\}$$
$$= c_3 E(n^{2r}\delta A^{2r}H^{-2r})E_1\{(H_n - H)^{2r}K^{-2r}\}.$$

By the Holder inequality, for $p>1$ and $q = (1 - 1/p)^{-1}$,

$$E_1\{(H_n - H)^{2r}K^{-2r}\} \le E_1^{1/p}(K^{-2rp})E_1^{1/q}(|H_n - H|^{2qr}).$$

Hence by Lemma B.1 and an application of the Marcinkiewicz–Zygmund inequality yields

$$E|T_{n,23}|^r \le c_4 n^{2r}\int H^{-2r}A^{2r}(nH)^{-2r}(n^{-qr}H)^{1/q}d\tilde{\tilde{H}}$$

which implies

$$\text{(B.4)} \qquad E|T_{n,23}|^r = O(n^{-r}).$$

Similarly for

$$T_{n,24} = -2\Sigma(K^{-2}A^2H^{-2}(H_n - H))(Z_i)\cdot\delta_i,$$

and

$$T_{n,25} = -(2n^{-1})\Sigma(n^2K^{-2}A^2H^{-1}(H_n - H))(Z_i)\cdot\delta_i,$$

it can be shown that

$$\text{(B.5)} \qquad E|T_{n,24}|^r = O(n^{-3r/2})$$

and

$$\text{(B.6)} \qquad E|T_{n,25}|^{2r} = O(n^{-r}).$$

From (B.2) through (B.6) we conclude that

$$\text{(B.7)} \qquad E|T_{n,2}|^{2r} = O(n^{-r}).$$

Now we consider the term $T_{n,1} = \int n^2K^{-2}(A_n^2 - A^2)d\tilde{\tilde{H}}_n$. Note that by Cauchy–Schwarz inequality we have

$$A_n^2 - A^2 = (A_n - A)^2 + 2A(A_n - A)$$
$$= \left(\int_\cdot^\infty F^{-1/2}(\hat{F}_n - F)F^{1/2}d\psi\right)^2 + 2A\left(\int_\cdot^\infty(\hat{F}_n - F)d\psi\right)$$
$$\le \left(\int_\cdot^\infty F^{-2}(\hat{F}_n - F)^2Fd\psi\right)\left(\int_\cdot^\infty Fd\psi\right) + 2A\left(\int_\cdot^\infty(\hat{F}_n - F)d\psi\right).$$

Therefore

$$|T_{n,1}| \le |\int n^2AK^{-2}(\int_{\cdot}^{\infty} F^{-2}(\hat{F}_n - F)^2 F d\psi) d\overset{\approx}{H}_n| + 2|\int n^2AK^{-2}(\int_{\cdot}^{\infty}(\hat{F}_n - F)d\psi)d\overset{\approx}{H}_n|.$$

Integration by parts on last two integrals yields

$$|T_{n,1}| \le c_5|\int(\int_0^{\cdot} n^2AK^{-2}d\overset{\approx}{H}_n)F^{-2}(\hat{F}_n - F)^2Fd\psi| + 2c_6|\int(\int_0^{\cdot} n^2AK^{-2}d\overset{\approx}{H}_n)(\hat{F}_n - F)d\psi|$$

$$= c_5|T_{n,11}| + 2c_6|T_{n,12}|, \text{ say.}$$

Since $T_{n,11}$ and $T_{n,12}$ are very similar, we just show that

$$E|T_{n,12}|^{2r} = O(n^{-r}).$$

Note that

$$|T_{n,12}| \le E^{(n)}\{|F^{-1}(\hat{F}_n - F)|\int_0^{\cdot} n^2AK^{-2}d\overset{\approx}{H}_n\}$$

where $E^{(n)}$ is the integral on $[0,\infty)$ with respect to $Fd\psi$.

Then by the Holder inequality, for $r>1$,

$$E^{(n)}\{|F^{-1}(\hat{F}_n - F)|\int_0^{\cdot} n^2AK^{-2}d\overset{\approx}{H}_n\}$$

$$\le (E^{(n)}\{|F^{-1}(\hat{F}_n - F)|\int_0^{\cdot} n^2AK^{-2}d\overset{\approx}{H}_n\}^r)^{1/r}\{E^{(n)}(1)\}^{1-1/r}$$

$$\le c_7E^{(n)}\{|F^{-1}(\hat{F}_n - F)|^r\int_0^{\cdot} n^{2r}A^rK^{-2r}d\overset{\approx}{H}_n\}.$$

Hence

$$E|T_{n,12}^{2r}| \le c_7E_*^{(n)}\{|F^{-1}(\hat{F}_n - F)|^{2r}\int_0^{\cdot} n^{4r}A^{2r}K^{-4r}d\overset{\approx}{H}_n\}$$

where $E_*^{(n)} = E \otimes E^{(n)}$. Let $p^{-1} + q^{-1} = 1$ and $p>1$. By the Holder inequality

$$E_*^{(n)}\{|F^{-1}(\hat{F}_n - F)|^{2r}\int_0^{\cdot} n^{4r}A^{2r}K^{-4r}d\overset{\approx}{H}_n\}$$

$$\le \{E_*^{(n)}|F^{-1}(\hat{F}_n - F)|^{2rp}\}^{1/p}\{E_*^{(n)}\{\int_0^{\cdot} n^{4r}A^{2r}K^{-4r}d\overset{\approx}{H}_n\}^q\}^{1/q}.$$

We shall show that the second term in the R.H.S. of the last inequality is finite and the first term is of order $O(n^{-r})$, $r>1$. Recall the representation

(1.1.5) and Lemma 1.1. First we show that $E_*^{(n)}\{\int_0^{\cdot} n^{4r}A^{2r}K^{-4r}d\tilde{\tilde{H}}_n\}^q<\infty$.

$$\begin{aligned} E_*^{(n)}\{\int_0^{\cdot} n^{4r}A^{2r}K^{-4r}d\tilde{\tilde{H}}_n\}^q &\le c_8E_*^{(n)}\{\int_0^{\cdot} n^{4qr}A^{2qr}K^{-4qr}d\tilde{\tilde{H}}_n\} \\ &\le c_9E\{\int_0^{\cdot} n^{4qr}A^{2qr}K^{-4qr}d\tilde{\tilde{H}}_n\}A(0) \\ &= c_9n^{4rq}E\{n^{-1}\Sigma\delta_iA^{2rq}(Z_i)K^{-4rq}(Z_i)\} \\ &= c_9n^{4rq}E\{\delta A^{2rq}(Z)E_1K^{-4rq}(Z)\} \\ &\le c_{10}\int A^{2rq}H^{-4rq}d\tilde{\tilde{H}} < \infty. \end{aligned}$$

Now by the Marcinkiewicz–Zygmund inequality and (1.1.7), it follows that

$$E_*^{(n)}|F^{-1}(\hat{F}_n - F)|^{2rp} = O(n^{-pr}).$$

Since p>1,

$$E|T_{n,12}^{2r}| = O(n^{-r}).$$

Similarly

$$E|T_{n,11}^{2r}| = O(n^{-r}).$$

Therefore

(B.8) $$E|T_{n,1}^{2r}| = O(n^{-r}).$$

Finally, note that $T_{n,3} = \int A^2H^{-2}d(\tilde{\tilde{H}} - \tilde{\tilde{H}})$ is an average of n iid mean zero bounded random variables. Thus an application of the Marcinkiewicz–Zygmund inequality yields

(B.9) $$E|T_{n,3}^{2r}| = O(n^{-r}),\ r>1.$$

Therefore (B.7), (B.8), (B.9) and Markov's inequality imply (1.1.13). □

REFERENCES

[1] Breslow, N. and Crowley, J. (1974). A large sample study of the life table and product–limit estimator under random censorship. Ann. Statist. **2**, 437–453.

[2] Carroll, R. J. (1978). On almost sure expansion for M–estimates. Ann. Statist. **6**, 314–318.

[3] Cheng, K. F. (1984). On almost sure representation for quantiles of the product–limit estimator with application. Sankhya, Ser. A, **46**, 426–443.

[4] Chow, Y. S. and Robbins, H. E. (1965). On the asymptotic theory of fixed width confidence interval for the mean. Ann. Math. Statist. **36**, 457–462.

[5] Chow, Y. S. and Teicher, H. (1978). *Probability Theory: Independence, Interchangeability, Martingales*, Springer–Verlag, New York.

[6] Chow, Y. S. and Yu, K. F. (1981). The performance of a sequential procedure for the estimation of the mean. Ann. Statist. **9**, 189–198.

[7] Főldes, A. and Rejtő, L. (1981). A LIL type result for the product–limit estimator. Z. Wahr. Verw. Gebiete **56**, 75–86.

[8] Gardiner, J. C. (1982). The asymptotic distribution of mortality rates in competing risk analyses. Scand. J. Statist. **9**, 31–36.

[9] Gardiner, J. C. and Susarla, V. (1983). Sequential estimation of the mean survival time under random censorship. Commun. Statist.– Sequential Anal. **2**(3), 201–223.

[10] Gardiner, J. C, Susarla, V. and van Ryzin, J. (1985a). Estimation of the median survival time under random censorship. *Adaptive Statistical Procedures and Related Topics*, **8**, 350–364.

[11] Gardiner, J. C, Susarla, V. and van Ryzin, J. (1985b). Fixed width confidence interval for the exponential mean survival time under random censorship. J. Statist. Plan. Infer. **12**, 153–159.

[12] Gill, R. (1983). Large sample behaviour of the product–limit estimator on the whole line. Ann. Statist. **11**, 49–58.

[13] Ghosh, M. and Mukhopadhyay, N. (1979). Sequential point estimation of the mean when the distribution is unspecified. Commun. Statist.–Theor. Meth. **A8**, 637–652.

[14] Hoem, J. M. (1976). The statistical theory of demographic rates. Scand. J. Statist. **3**, 169–185.

[15] Hoem, J. M. (1987). Statistical analysis of a multiplicative model and its application to the standardization of vital rates: A review. Int. Statist. Rev. **55** (2), 119–152.

[16] Huber, P. J. (1964). Robust estimation of a location parameter. Ann. Math. Statist. **35**, 73–101.

[17] Huber, P. J. (1981). *Robust Statistics.* Wiley & Sons, New York.

[18] Jureckova, J and Sen, P. K. (1981). Sequential procedures based on M–estimators with discontinuous score functions. J. Statist. Plan.Infer. **5**, 253–266.

[19] Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. J. Amer. Statist. Assoc. **53**, 457–481.

[20] Koul, H., Susarla, V. and van Ryzin, J. (1981). Regression analysis with randomly right–censored data. Ann. Statist. **9**, 1276–1288.

[21] Lo, S. H. and Singh, K. (1986). The product–limit estimator and the bootstrap (some asymptotic representations). Prob. Theo. Relat. Fields. **71**, 455–465.

[22] Millar, P.W. (1985). Non–parametric applications of an infinite dimensional convolution theorem. Z. Wahr. Verw. Gebiete **68**, 545–556.

[23] Reid, N. (1981). Influence functions for censored data. Ann. Statist. **9**, 78–92.

[24] Robbins, H. (1959). Sequential estimation of the mean of the normal population. *Probability and Statistics* (Harold Cramer Volume). Almqvist and Wiksell, Uppsala, Sweden. 235–245.

[25] Schick, A., Susarla, V. and Koul, H. (1987). Efficient estimation of functionals with censored data. To appear in Statist. Decision.

[26] Sen, P.K. (1980). On nonparametric sequential point estimation of location based on general rank order statistics. Sankhya, Ser. A **42**, 201–218.

[27] Sen, P.K.(1981). *Sequential Nonparametrics,* Wiley & Sons, New York.

[28] Sen, P. K. and Ghosh, M. (1981). Sequential point estimation of estimable parameters based on U–statistics. Sankhya, ser. A **43**, 331–344.

[29] Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics,* Wiley & Sons, New York.

[30] Shorack, G.R. and Wellner, J.A. (1986). *Empirical Processes with Applications to Statistics.* Wiley & Sons, New York

[31] Sriram, T. N. (1987). Sequential estimation of the mean of a first–order stationary autoregressive process. Ann. Statist. **15**, 1079–1090.

[32] Starr, N. (1966). On the asymptotic efficiency of sequential procedure for estimating the mean. Ann. Math. Statist. **37**, 1173–1185.

[33] Starr, N. and Woodroofe, M. (1969). Remarks on sequential point estimation. Proc. Nat. Acad. Sci. U.S.A. **63**, 285–288.

[34] Starr, N. and Woodroofe, M. (1972). Further remarks on sequential point estimation: The exponential case.
Ann. Math. Statist. **43**, 1147–1154.

[35] Susarla, V. and van Ryzin, J. (1980). Large sample theory for an estimator of the mean survival time from censored samples. Ann. Statist. **8**, 1002–1016.

[36] Wellner, J. A. (1982). Asymptotic optimality of the product–limit estimator. Ann. Statist. **10**, 595–602.

[37] Woodroofe, M. (1982). *Nonlinear Renewal Theory in Sequential Analysis*, CBMS–NSF Regional Conference Series in Applied Mathematics. SIAM.