WEIGHTED L^2 QUANTILE DISTANCE ESTIMATORS FOR RANDOMLY CENSORED DATA

by

R. L. Eubank and V.N. LaRiccia

Technical Report No. 149
Department of Statistics ONR Contract

November, 1981

Research sponsored by the Office of Naval Research
Contract N00014-75-C-0439

DEPARTMENT OF STATISTICS
Southern Methodist University
Dallas, Texas 75275

Weighted L^2 Quantile Distance Estimators For Randomly Censored Data

by

R. L. Eubank and V. N. LaRiccia
Southern Methodist University and University of Nebraska-Lincoln

Short Title: Estimators for Censored Data

Summary: The asymptotic properties of a family of minimum quantile function distance estimators for randomly censored data sets are considered. These procedures produce an estimator of the parameter vector that minimizes a weighted L^2 distance measure between the Kaplan-Meier quantile function and an assumed parametric family of quantile functions. Regularity conditions are provided which insure that these estimators are consistent and asymptotically normal. An optimal weight function is derived for single parameter families, which, for location/scale families, results in censored sample analogs of estimators such as those suggested by Parzen (1979a, 1979b), and Weiss and Wolfowitz (1970).

1. Introduction. In this paper we consider the problem of parameter estimation from randomly censored data sets. A general method of estimation is presented for cases when the data is assumed to be from a known parametric family. The technique is based on the minimization of a weighted L^2 distance measure between the Kaplan-Meier empirical quantile function and the assumed parametric family of quantile functions and is applicable to most common distributions.

Let $X_1, \ldots, X_n$ denote the true survival times of n individuals which are assumed to be a random sample from the distribution function (d.f.) $F(x;\underline{\theta}^o)$ where $\underline{\theta}^o$ is a fixed, possibly unknown element of a known set or region $\Theta \subset \mathbb{R}^s$. Further let $Y_1,\ldots,Y_n$ denote n independent identically distributed censoring random variables with common distribution function H that are also assumed to be independent of the X_i's. In the random censoring model one observes not the X_i's but, instead, the pairs of random variables (Z_i,δ_i) where $Z_i = \min(X_i,Y_i)$ and $\delta_i = I_{\{X_i \leq Y_i\}}$ with I denoting the indicator function. The d.f. of the Z_i's, F^*, is then given by the relation

$$(1.1) \qquad 1-F^*(x;\underline{\theta}^o) = [1-F(x;\underline{\theta}^o)][1-H(x)].$$

An important problem associated with this model is the estimation of the parameter vector, $\underline{\theta}^o$, from the observed data.

Let $F_n(x)$ denote the Kaplan-Meier estimator of the d.f. $F(x;\underline{\theta}^o)$ (Kaplan and Meier (1958)) with associated empirical quantile function defined by

$$(1.2) \qquad Q_n(u) = \inf_x\{F_n(x) \geq u\} .$$

Estimation problems pertaining to specific forms for the vector $\underline{\theta}^o$ have been addressed by Sander (1975a, 1975b, 1975c) and Susarla and Van Ryzin (1980) and Reid (1981) using estimators based, explicity, on both F_n and Q_n. In contrast, we develop an estimation procedure applicable to general $\underline{\theta}^o$ under the assumption that the functional form of F is known. Thus when, for instance, $\underline{\theta}^o$ consists of only a location and scale parameter, the model we assume is the censored sample analog of the classical location and scale parameter model.

Minimum distance estimation procedures based on the d.f., the probability density function, the characteristic function, and the quantile function have been proposed and both their large and small sample properties, for non-randomly censored data sets, have been extensively investigated. Specific minimum distance procedures have been shown to possess excellent robustness properties, to be consistent, asymptotically normal, and in some cases fully efficient. While a detailed discussion of these points is beyond the scope of this paper, the interested reader is referred to Beran (1977), Parr and Schucany (1980), and Millar (1981) for discussions of techniques which utilize the d.f.. Estimation procedures formulated in the quantile function domain may be found, for the case of location and/or scale parameter estimation, in Parzen (1979a) and Eubank (1981) and, for more general parametrizations, in LaRiccia and Wehrly (1981) and LaRiccia (1981). We note that certain results presented in subsequent sections may be regarded as censored sample analogs of techniques developed in Parzen (1979a) and LaRiccia (1981) and are obtained through the use of work by Sander (1975a, 1975b) and Reid (1981) on the convergence of the empirical quantile function and linear functions of order statistics for randomly censored data.

In Section 2 we define our estimator and present our principal results regarding its asymptotic properties. The proofs are provided in Section 3. Finally, in Section 4, we discuss the estimator's efficiency and robustness properties and provide an optimal weight function for single parameter families.

2. <u>A family of minimum quantile function distance estimators</u>. For each $\underline{\theta} \in \Theta$ define the quantile function associated with $F(x;\underline{\theta})$

by

$$(2.1) \qquad Q(u;\underline{\theta}) = \inf\{x : F(x;\underline{\theta}) \geq u\} \ .$$

For a given weight function $W(u;\underline{\theta})$ mapping $(0,1) \times \Theta \to \mathbb{R}$ and associated real valued functions $W_{\beta_1}(\underline{\theta})$ and $W_{\beta_2}(\underline{\theta})$, the <u>minimum quantile function distance estimator of the true parameter</u>, $\underline{\theta}^{o}$, is any vector which minimizes

$$(2.2) \qquad R(Q_n,\underline{\theta}) = \int_{\beta_1}^{\beta_2} W(u;\underline{\theta})[Q_n(u) - Q(u;\underline{\theta})]^2 du$$

$$+ \ W_{\beta_1}(\underline{\theta})[Q_n(\beta_1) - Q(\beta_1;\underline{\theta})]^2 + W_{\beta_2}(\underline{\theta})[Q_n(\beta_2) - Q(\beta_2;\underline{\theta})]^2$$

over all $\underline{\theta} \in \Theta$ where β_1 and β_2 are fixed real numbers satisfying $0 \leq \beta_1 < \beta_2 \leq 1$ but are otherwise arbitrary. Since individual members of this family of estimators are distinguished by their specific weight function we adopt the notation $\hat{\underline{\theta}}(Q_n,W)$ for this estimator. For notational convenience it will be useful to have an expression which incorporates all three of the functions W, W_{β_1} and W_{β_2}. Therefore, for any function $Z(u;\underline{\theta})$ we define

$$(2.3) \qquad \int_{\beta_1}^{\beta_2} W^*(u;\underline{\theta})Z(u;\underline{\theta})du = \int_{\beta_2}^{\beta_1} W(u;\underline{\theta})Z(u;\underline{\theta})du$$

$$+ \sum_{i=1}^{2} W_{\beta_i}(\underline{\theta})Z(\beta_i;\underline{\theta})$$

and (2.2) may now be written as

$$(2.4) \qquad R(Q_n;\underline{\theta}) = \int_{\beta_1}^{\beta_2} W^*(u;\underline{\theta})[Q_n(u) - Q(u;\underline{\theta})]^2 du \ .$$

Weight function selection strategies will be discussed in Section 4.

In general, the computation of $\underline{\hat{\theta}}(Q_n,W)$ can be accomplished using standard iterative techniques. Estimator computation is particularly simple for location and scale parameter families when the weight function is independent of $\underline{\theta}$ as, in this instance, the estimator is readily seen to have a closed form.

One advantage of the minimum distance estimation technique is that $R(Q_n,\underline{\hat{\theta}}(Q_n,W))$ provides a measure of the goodness-of-fit of the assumed parametric family. In addition, the quantile function approach has the consequence that the estimation procedure is directly related to Q-Q plotting techniques. Therefore, the estimators can be easily incorporated into a statistical package which not only estimates the parameters but also provides checks for the appropriateness of the assumed parametric family and graphs which can be employed to, perhaps, suggest a more suitable family of distributions.

The asymptotic behaviour and distribution theory for $\underline{\hat{\theta}}(Q_n,W)$ is the subject of Theorem 1. To prove these results we require certain restrictions on both W and Q. Therefore, let G denote the class of left continuous functions on (0,1) that are of bounded variation on $(\gamma,1-\gamma)$ for all $0 < \gamma < \frac{1}{2}$ and, for h, g ε G define

$$d(g,h) = \sup_{x\varepsilon[\beta_1,\beta_2]} |h(x)-g(x)| \quad .$$

Also for any function $Z(u;\underline{\theta})$ let $Z^i(u;\underline{\theta}) = \frac{\partial Z(u;\underline{\theta})}{\partial\theta_i}$,

$Z^{ij}(u;\underline{\theta}) = \frac{\partial^2 Z(u;\underline{\theta})}{\partial\theta_i\partial\theta_j}$, $Z'(u;\underline{\theta}) = \frac{\partial Z(u;\underline{\theta})}{\partial u}$ and $Z_{\underline{\theta}} \doteqdot Z(\cdot;\underline{\theta})$.

The following assumptions are then required for Theorem 1:

(1) $d(Q_n, Q_{\underline{\theta}^o}) \to_{\underline{P}} 0$, where $\to_{\underline{P}}$ denotes convergence in probability.

(2) $\sqrt{n}\, d(Q_n, Q_{\underline{\theta}^o}) = O_{\underline{P}}(1)$, where $O_{\underline{P}}$ denotes probability order.

(3) Let $\underline{V}_n$ denote the vector having ith component

$$\sqrt{n}\{\int_{\beta_1}^{\beta_2} W^*(u;\underline{\theta}^o) Q^i(u;\underline{\theta}^o)(Q_n(u) - Q(u;\underline{\theta}^o))du\}$$

where the integral is defined as in (2.2). Then $\underline{V}_n$ converges in law to the $N_s(\underline{0}, A)$ distribution, denoted $\underline{V}_n \to_L N_s(\underline{0}, A)$, where the ijth element of A is

$$a_{ij} = \int_{\beta_1}^{\beta_2}\int_{\beta_1}^{\beta_2} W^*(u;\underline{\theta}^o) W^*(v;\underline{\theta}^o) \frac{Q^i(u;\underline{\theta}^o) Q^j(v;\underline{\theta}^o)}{fQ(u;\underline{\theta}^o) fQ(v;\underline{\theta}^o)} K(u,v)dudv \tag{2.5}$$

with

$$K(u,v) = (1-u)(1-v)\int_0^{\min(u,v)} [(1-s)^2(1-HQ(s;\underline{\theta}^o))]^{-1} ds \quad , \tag{2.6}$$

$HQ(u;\underline{\theta}^o) = H(Q(u;\underline{\theta}^o))$, and $fQ(u;\underline{\theta}^o) = F'(Q(u;\underline{\theta}^o);\ \underline{\theta}^o)$ is the density-quantile function corresponding to F (c.f. Parzen (1979a)).

(4) For fixed $\underline{\theta}$, the functions $W(u;\underline{\theta})$, $W^i(u;\underline{\theta})$, $W^{ij}(u;\underline{\theta})$, $Q^i(u;\underline{\theta})$ and $Q^{ij}(u;\underline{\theta})$ are continuous on $[\beta_1, \beta_2]$ and, for fixed u, in a neighborhood of $\underline{\theta}^o$. $W^i_{\beta_k}(\underline{\theta})$, $W^{ij}_{\beta_k}(\underline{\theta})$, $k = 1,2$, are continuous in $\underline{\theta}$.

(5) The functions $W^i(u;\underline{\theta})$, $W(u;\underline{\theta})Q^i(u;\underline{\theta})$, $W^{ij}(u;\underline{\theta})$, $W^i(u;\underline{\theta})$, $Q^j(u;\underline{\theta})$ and $W(u;\underline{\theta})Q^{ij}(u;\underline{\theta})$ are bounded by integrable functions uniformly for all $\underline{\theta}$ in the neighborhood of $\underline{\theta}^o$.

(6) The sxs matrix $B(Q_{\underline{\theta}^o},\underline{\theta}^o)$ having ij^{th} element

$$b_{ij} = \int_{\beta_1}^{\beta_2} W^*(u;\underline{\theta}^o)Q^i(u;\underline{\theta}^o)Q^j(u;\underline{\theta}^o)\,du \tag{2.7}$$

is positive definite.

Theorem 1 Under the regularity conditions (1) - (6)

(i) As $n\to\infty$ there exists, with probability tending to one, a unique function, $\hat{\underline{\theta}}(Q_n,W)$, which locally minimizes (2.1),

(ii) $\hat{\underline{\theta}}(Q_n,W)$ is a consistent estimator of $\underline{\theta}^o$,

(iii) $\sqrt{n}(\hat{\underline{\theta}}(Q_n,W) - \underline{\theta}^o) \to N_s(0,C)$ where

$$C = [B(Q_{\underline{\theta}^o},\underline{\theta}^o)]^{-1}A[B(Q_{\underline{\theta}^o},\underline{\theta}^o)]^{-1} \tag{2.8}$$

with A and B as defined in (2.5) and (2.7).

Conditions (1) - (6) can be replaced by many different sets of restrictions. In particular conditions which imply (1) and (2) can be derived from the results of Sander (1975a) whereas restrictions implying (3) can be found in Sander (1975b) and Reid (1981). As an illustration of alternative conditions we present the following corollary.

Corollary . Let assumptions (4) - (6) be satisfied and further assume that

(i) $0 < \beta_1 < \beta_2 < 1$ and H is a continuous function satisfying $HQ(\beta_2;\underline{\theta}^o) < 1$,

(ii) F is a strictly increasing function with $F(0;\underline{\theta}^o) = 0$,

(iii) $W(u;\underline{\theta}^o)Q^i(u;\underline{\theta}^o)$ is differentiable.

Then, the conclusions of Theorem 1 hold.

Before proving Theorem 1 in the next section we note that an alternative approach to our method of proof could be taken wherein $\hat{\underline{\theta}}(Q_n,W)$ is treated as a functional on the metric space (G,d). Thus, assumptions similar to those in Beran (1977) on the differentiability of the functional $\hat{\underline{\theta}}(Q_n,W)$ could be used to replace conditions (4)-(6) although the present conditions seem easier to verify in practice. However, if $W(u;\underline{\theta}^o) > 0$ for all $u \in [\beta_1,\beta_2]$ it is usually easier to check the condition $\sum_{i=1}^{s} a_i Q^i(u;\underline{\theta}^o) \neq 0$ for all $\underline{a} = (a_1,\ldots,a_s)^t \neq 0$ which, in this case, is equivalent to (6). We note that this latter condition is an assumption on the identifiability of the parameters.

3. Proofs. We now prove Theorem 1. Let $D(Q_n,\underline{\theta}) = \dfrac{\partial R(Q_n,\underline{\theta})}{\partial \underline{\theta}}$ and $J(Q_n,\underline{\theta}) = \dfrac{\partial^2 R(Q_n,\underline{\theta})}{\partial \underline{\theta}^2}$. Then, D is the s x 1 vector with i^{th} element

$$R^i(Q_n;\underline{\theta}) = \int_{\beta_1}^{\beta_2}\{(Q_n(u)-Q(u;\underline{\theta}))^2 W^{*i}(u;\underline{\theta}) - 2W^*(u;\underline{\theta})Q^i(u;\underline{\theta})(Q_n(u)-Q(u;\underline{\theta}))\}du$$

where the W^{*i} notation and the integrals are defined in a manner analogous to (2.2). Similarly, J is the s x s Jacobian matrix with ij^{th} element

$$R^{ij}(Q_n;\underline{\theta}) = \int_{\beta_1}^{\beta_2}\{(Q_n(u)-Q(u;\underline{\theta}))^2 W^{*ij}(u;\underline{\theta})-2(Q_n(u)-Q(u;\underline{\theta}))[W^{*i}(u;\underline{\theta})Q^j(u;\underline{\theta})$$
$$+ W^{*j}(u;\underline{\theta})Q^i(u;\underline{\theta})+W^*(u;\underline{\theta})Q^{ij}(u;\underline{\theta})]+2W^*(u;\underline{\theta})Q^i(u;\underline{\theta})Q^j(u;\underline{\theta})\}du .$$

Now, employing the compactness of $[\beta_1,\beta_2]$ and conditions (4) - (5) tedious but straightforward calculations give the following three results:

(A) In an open neighborhood of $(Q_{\underline{\theta}^o},\underline{\theta}^o)$, $R^i(Q_n,\underline{\theta})$ and $R^{ij}(Q_n,\underline{\theta})$ are continuous functions with respect to the metric ρ, defined for g, $h \in G$ and $\underline{\theta},\underline{\alpha} \in \Theta$ by

$$\rho[(h,\underline{\theta}),(g,\underline{\alpha})] = \max\{d(g,h),\ |\theta_1-\alpha_1|,\ldots,|\theta_s-\alpha_s|\} ,$$

(B) $\sqrt{n}\int_{\beta_1}^{\beta_2}(Q_n(u)-Q(u;\underline{\theta}^o))^2 W^{*i}(u;\underline{\theta}^o)du \to_p 0$ and

(C) $R^{ij}(Q_n,\hat{\underline{\theta}}) \to R^{ij}(Q_{\underline{\theta}^o},\underline{\theta}^o)$ if $\hat{\underline{\theta}} \to_p \underline{\theta}^o$.

Since $D(Q_{\underline{\theta}^o},\underline{\theta}^o) = \underline{0}$ and $J(Q_{\underline{\theta}^o},\underline{\theta}^o)$ is positive definite it follows, using (A), that the conditions of the implicit function theorem are satisfied. Consequently, there exists an open neighborhood of $(Q_{\underline{\theta}^o},\underline{\theta}^o)$ and a unique continuous mapping, $\hat{\underline{\theta}}(Q_n,W)$, of an open ball, $\mathcal{B}$, containing $Q_{\underline{\theta}^o}$, into $\mathbb{R}^s$ such that $d(Q_n, \hat{\underline{\theta}}(Q_n,W)) = 0$ and $\hat{\underline{\theta}}(Q_{\underline{\theta}^o},W) = \underline{\theta}^o$. Further, as a result of (1), with probability tending to one, Q_n is in $\mathcal{B}$ so that there exists, also with probability tending to one, a unique point $\hat{\underline{\theta}}(Q_n,W)$ such that $D(Q_n,\hat{\underline{\theta}}(Q_n,W))= 0$. To see that $\hat{\underline{\theta}}(Q_n,W)$ minimizes $R(Q_n,\underline{\theta})$ note that, using (A), $J(Q_n,\hat{\underline{\theta}}(Q_n,W)) \to_p B(Q_{\underline{\theta}^o}, \underline{\theta}^o)$ which is positive definite by assumption (6). Thus, result (i) has been shown. Result (ii), the consistency of $\hat{\underline{\theta}}(Q_n,W)$, is an immediate consequence of the continuity of $\hat{\underline{\theta}}(\cdot,W)$ and the fact that $\hat{\underline{\theta}}(Q_{\underline{\theta}^o},W) = \underline{\theta}^o$.

By the mean value theorem

$$D(Q_n,\underline{\theta}^o) = D(Q_n,\hat{\underline{\theta}}(Q_n,W)) + J(Q_n,\overline{\underline{\theta}})(\underline{\theta}^o - \hat{\underline{\theta}}(Q_n,W))$$

where $\overline{\underline{\theta}} = \underline{\theta}^o + \Delta(\underline{\theta}^o - \hat{\underline{\theta}}(Q_n,W))$ for some sxs diagonal matrix, Δ, with all elements between 0 and 1. Since $D(Q_n,\hat{\underline{\theta}}(Q_n,W)) = 0$,

$$\sqrt{n}[\hat{\underline{\theta}}(Q_n,W)-\underline{\theta}^o] = -J(Q_n,\overline{\underline{\theta}})^{-1}\sqrt{n}D(Q_n,\underline{\theta}^o) .$$

Result (iii) now follows using (B), (C), assumption (3) and a Slutsky type argument.

To verify the corollary we first note that it is shown in Sander (1975a, Corollary 1) that $\sqrt{n}[Q_n(u)-Q(u;\underline{\theta}^o)]$ converges weakly to $(-1/fQ(u;\theta))X(u)$ on $[\beta_1,\beta_2]$ where $X(\cdot)$ is a zero mean Gaussian process with covariance kernel (2.5). Thus, it can be readily shown that assumptions (1) and (2) are satisfied. This fact also implies that, for $i=1,\ldots,s$,

$$\sum_{k=1}^{2} W_{\beta_k}(\underline{\theta}^o)Q^i(\beta_k;\underline{\theta}^o)(Q_n(\beta_k) - Q(\beta_k;\underline{\theta}^o)) \to_L N(0,\sigma_1^2)$$

where

$$\sigma_1^2 = \sum_{\ell=1}^{2}\sum_{k=1}^{2} W_{\beta_\ell}(\underline{\theta}^o)Q^i(\beta_\ell,\underline{\theta}^o)W_{\beta_k}(\underline{\theta}^o)Q^i(\beta_k;\underline{\theta}^o)[fQ(\beta_\ell;\underline{\theta}^o)fQ(\beta_k;\underline{\theta}^o)]^{-1}K(\beta_\ell,\beta_k) .$$

Since $W(u;\underline{\theta}^o)Q^i(u;\underline{\theta}^o)$ is differentiable on $[\beta_1,\beta_2]$ the results of Example 3 of Reid (1981) have the consequence that

$$\int_{\beta_1}^{\beta_2} W(u;\underline{\theta}^o)Q^i(u;\underline{\theta}^o)(Q_n(u)-Q(u;\underline{\theta}^o))du \to_L N(0,\sigma_2^2)$$

where

$$\sigma_2^2 = \int_{\beta_1}^{\beta_2}\int_{\beta_1}^{\beta_2} W(v;\underline{\theta}^o)Q^i(v;\underline{\theta}^o)W(u;\underline{\theta}^o)Q^i(u;\underline{\theta}^o)[fQ(v;\underline{\theta}^o)fQ(u;\underline{\theta}^o)]^{-1}K(v,u)dvdu .$$

Then, an argument similar to that used in Shorack (1972) gives (3) and the conditions of Theorem (2.1) are satisfied.

4. Robustness and efficiency. In this section we discuss the robustness and efficiency properties of the estimator $\hat{\underline{\theta}}(Q_n,W)$. Using an approach similar to that of Parzen (1979a) an optimal weight function is provided for estimation in single parameter families. In the case of location or scale families with no censoring (H=0) the resulting estimators agree with estimators proposed by Parzen (1979a, 1979b) and, consequently, are similar to those given by Weiss and Wolfowitz (1970). The results in this section, therefore, provide an extension to randomly censored samples of the work of these authors.

It follows from the proof of Theorem 1 that, asymptotically, $\hat{\underline{\theta}}(Q_n,W)$ is a weighted sum of s linear functions of the order statistics.

Thus the influence curve for these estimators can be determined using the techniques developed in Reid (1981). For a given parametric family of distributions, weight functions with specific properties or which provide specific types of protection can then be derived. In particular, by proper selection of $W(u;\underline{\theta})$ an estimator which is locally robust or one that has specified efficiency properties can be obtained.

The remainder of this section is devoted to the problem of weight function selection. To obtain an optimal weight function we first note, as in Parzen (1979a), that the covariance kernel, K, in (2.6) generates a reproducing kernel Hilbert space, $\mathcal{H}(K)$, which provides us with a natural measure of information on $[\beta_1,\beta_2]$. The properties of $\mathcal{H}(K)$ that impact on our present objective will now be briefly discussed. The reader is referred to Aronszajn (1950) and Parzen (1961a, 1961b) for a more detailed presentation of, respectively, the theory of reproducing kernels and their role in inference for stochastic processes.

Assuming that H admits a continuous density, h, it follows from Sacks and Ylvisaker (1966) that $\mathcal{H}(K)$ consists of continuous functions on $[\beta_1,\beta_2]$ having finite $\mathcal{H}(K)$ norm where, for $g\in\mathcal{H}(K)$, the norm is given by

$$\begin{aligned}||g||_K^2 = &\int_{\beta_1}^{\beta_2}[g'(u)]^2[1-HQ(u;\underline{\theta})]du \\ &+\int_{\beta_1}^{\beta_2}\left[\frac{g(u)}{1-u}\right]^2 (1-u)\ \frac{hQ(u;\underline{\theta})}{fQ(u;\underline{\theta})}\,du \\ &+ g(\beta_1)^2\left[\frac{1}{K(\beta_1,\beta_1)} - \frac{1-HQ(\beta_1;\underline{\theta})}{1-\beta_1}\right] \\ &+\frac{g(\beta_2)^2}{1-\beta_2}\,[1-HQ(\beta_2;\underline{\theta})] \quad .\end{aligned} \tag{4.1}$$

The inner product in $\mathcal{H}(K)$ will be denoted by $\langle\cdot,\cdot\rangle_K$. When considered as a function of s (for fixed t)K(s,t) possess the so called reproducing property that for $g \varepsilon \mathcal{H}(K)$

(4.2) $\langle g, K(\cdot,t)\rangle_K = g(t)$.

We now, and in subsequent discussions, impose the regularity condition that $fQ(u;\underline{\theta})Q^i(u;\underline{\theta}) \; \varepsilon \; \mathcal{H}(K)$ and define the information measure

(4.3) $I(\theta_i) = ||fQ(\cdot;\underline{\theta})Q^i(\cdot;\underline{\theta})||_K^2$.

To justify this definition we first note that the densities corresponding to the censored and uncensored observations are, respectively,

$$f^c(x;\underline{\theta}) = [1-F(x;\underline{\theta})]h(x)$$

and

$$f^u(x;\underline{\theta}) = [1-H(x)]f(x;\underline{\theta}) \; .$$

Then by making the change of variable $X = Q(u;\underline{\theta})$ it can be shown that, for instance, when $\beta_1 = 0$ and $\beta_2 = 1$, $I(\theta_i)$ becomes

(4.4)
$$I(\theta_i) = \int_0^\infty \left[\frac{f^i(x;\underline{\theta})}{f(x;\underline{\theta})}\right]^2 f^u(x;\underline{\theta})\,dx + \int_0^\infty \left[\frac{F^i(x;\underline{\theta})}{1-F(x;\underline{\theta})}\right]^2 f^c(x;\underline{\theta})\,dx$$

This is now recognized as the Fisher information corresponding to the parameter θ_i upon examination of the form of the likelihood for the Z_i's given in, for example, Kalbfleisch and Prentice (1980). Similarly, $I(\theta_i)$ can be seen to provide a measure of information in the Fisher sense when $0 < \beta_1 < \beta_2 < 1$.

Now suppose that there is only one unknown parameter, $\theta_1 = \theta$, and let

$$G(u;\theta) = (1-u)\int_0^u \frac{1}{(1-\omega)^2[1-HQ(\omega;\theta)]}\, d\omega \quad .$$

Under the assumption that $fQ(u;\theta)Q^1(u;\theta)$ is twice continuously differentiable define

$$(4.5) \quad \phi(u;\theta) = -[fQ(u;\theta)Q^1(u;\theta)]''[1-HQ(u;\theta)] - \left[\frac{f_0Q_0(u;\theta)Q^1(u;\theta)}{1-u}\right]'(1-u)\,\frac{hQ(u;\theta)}{fQ(u;\theta)}$$

for $\beta_1 < u < \beta_2$ with

$$(4.6) \quad \phi_{\beta_1}(\theta) = \frac{\{fQ(\beta_1;\theta)Q^1(\beta_1;\theta)G'(\beta_1;\theta) - G(\beta_1;\theta)(fQ\cdot Q^1)'(\beta_1;\theta)\}[1-HQ(\beta_1;\theta)]}{G(\beta_1;\theta)}$$

and

$$(4.7) \quad \phi_{\beta_2}(\theta) = \frac{\{fQ(\beta_2;\theta)Q^1(\beta_2;\theta) + (1-\beta_2)(fQ\cdot Q^1)'(\beta_2;\theta)\}[1-HQ(\beta_2;\theta)]}{1-\beta_2}$$

where $(fQ\cdot Q^1)'(\beta_k;\theta) = [fQ(u;\theta)Q^1(u;\theta)]'\Big|_{u=\beta_k}$. Then, using results from Sacks and Ylvisaker (1966) it follows that $fQ(u;\theta)Q^1(u;\theta)$ admits the representation

$$(4.8) \quad fQ(u;\theta)Q^1(u;\theta) = \int_{\beta_1}^{\beta_2} \phi(s;\theta)K(s,u)du + \sum_{k=1}^{2} \phi_\beta(\theta)K(\beta_k,u)$$

for $u \in [\beta_1,\beta_2]$. By using (4.8) in conjuction with the reproducing property for K it is possible to verify the two important identities

$$||fQ(\cdot;\theta)Q^1(\cdot;\theta)||_K^2 = \int_{\beta_1}^{\beta_2} fQ(u;\theta)Q^1(u;\theta)\phi(u;\theta)du$$

(4.9)
$$+ \sum_{k=1}^{2} \phi_{\beta_k}(\theta)fQ(\beta_k;\theta)Q^1(\beta_k;\theta)$$

and

$$||fQ(\cdot;\theta)Q^1(\cdot;\theta)||_K^2 = \int_{\beta_1}^{\beta_2}\int_{\beta_1}^{\beta_2}\phi(u;\theta)\phi(v;\theta)K(u,v)dudv$$

(4.10)
$$+ 2\sum_{k=1}^{2} \phi_{\beta_k}(\theta)\int_{\beta_1}^{\beta_2} \phi(u;\theta)K(\beta_k,u)du$$

$$+ \sum_{k=1}^{2}\sum_{\ell=1}^{2} \phi_{\beta_k}(\theta)\phi_{\beta_\ell}(\theta)K(\beta_k,\beta_\ell) \quad .$$

Now define the weight function

(4.11) $$W(u;\theta) = \frac{\phi(u;\theta)fQ(u;\theta)}{Q^1(u;\theta)} , \qquad \beta_1 < u < \beta_2$$

with

(4.12) $$W_{\beta_k}(\theta) = \frac{\phi_{\beta_k}(\theta)fQ(\beta_k;\theta)}{Q^1(\beta_k;\theta)} , \quad k = 1,2.$$

Using this weight function in (2.4) and (2.6) and applying the identities (4.9) - (4.10) it can be concluded, from Theorem 1, that

$$\sqrt{n}(\hat{\theta}(Q_n,W) - \theta^o) \to_L N(0,1/I(\theta^o))$$

and hence, the weight function (4.11) - (4.12) is optimal for the estimation of θ^o.

The presence of 1-HQ(u;θ) in the weight function has the consequence that the weight function depends on the censoring distribution which may not be known. Although the details are beyond the scope of this paper, such difficulties can be avoided by using strongly consistent estimators of H and h in W and W_{β_k}, k = 1,2.

To provide further insight into the properties of this weight function consider the special case when θ is a scale parameter and there is no censoring, i.e., H = 0. In this case, we have

$$F(x;\theta) = F(x/\theta)$$

$$f(x;\theta) = \frac{1}{\theta}f(x/\theta)$$

$$Q(u;\theta) = \theta Q(u)$$

$$fQ(u;\theta) = \frac{1}{\theta} fQ(u)$$

and

$$Q^1(u;\theta) = Q(u)$$

where F(x), Q(u) and fQ(u) are all specified functions. Equations (4.5) - (4.7) now become

$$(4.13) \qquad \phi(u) = -\{fQ(u)Q(u)\}''$$

$$(4.14) \qquad \phi_{\beta_1} = \frac{1}{\beta_1} fQ(\beta_1)Q(\beta_1)-(fQ\cdot Q)'(\beta_1)$$

and

$$(4.15) \qquad \phi_{\beta_2} = \frac{1}{1-\beta_2} fQ(\beta_2)Q(\beta_2) + (fQ\cdot Q)'(\beta_2)$$

so that W is independent of θ. Consequently, $\hat{\theta}(Q_n,W)$ may be expressed in the closed form

$$(4.16) \quad \hat{\theta}(Q_n,W) = \frac{\int_{\beta_1}^{\beta_2} \phi(u)fQ(u)Q_n(u)du + \sum_{k=1}^{2} \phi_{\beta_k} fQ(\beta_k)Q_n(\beta_k)}{\int_{\beta_1}^{\beta_2} \phi(u)fQ(u)Q(u)du + \sum_{k=1}^{2} \phi_{\beta_k} fQ(\beta_k)Q(\beta_k)} .$$

This is precisely the fully efficient estimator of θ considered by Parzen (1979a,1979b). Identical results hold for location parameter estimation as well. Parzen has noted that these estimators are similar to those presented in Weiss (1964) and Weiss and Wolfowitz (1970). Consequently, the minimum quantile function distance procedures developed in this paper, when incorporated with the weight functions derived in this section may be viewed as providing,for location/scale models, censored sample analogs of these type of estimators.

REFERENCES

1. Aronszajn, N.: Theory of reproducing kernels. Trans. Amer. Math. Soc. 68, 337-404 (1950).

2. Beran, Rudolf: Minimum Hellinger distance estimates for parametric models. Ann. Statist. 5, 445-463 (1977).

3. Eubank, R. L.: A density-quantile function approach to optimal spacing selection. Ann. Statist. 9, 494-500 (1981).

4. Kalbfleisch, J.D. and Prentice, R. L.: The statistical analysis of failure time data, John Wiley, New York. (1980).

5. Kaplan, E.L. and Meier, P.: Nonparametric estimation from incomplete observations. J. Amer. Statist. Assoc. 53, 457-581 (1958).

6. LaRiccia, V.N.: Asymptotic properties of weighted L^2 quantile distance estimators. To appear. (1981).

7. LaRiccia, V.N. and Wehrly, T.E.: Asymptotic properties of a family of minimum quantile distance estimators. To appear. (1981).

8. Millar, P.: Robust estimation via minimum distance methods. Z. Wahrscheinlichkeitstheorie verw. Geb. 55, 73-89 (1981).

9. Parr, W.C. and Schucany, W.R.: Minimum distance and robust estimation. J.Amer. Statist. Assoc. 75, 616-624 (1980).

10. Parzen, E.: An approach to time series analysis. Ann. Math. Statist. 32, 951-989 (1961a).

11. Parzen, E.: Regression analysis for continuous parameter time series. Proc. 4th Berkeley Sympos. Math., Statist. and Prob., Vol. I. Univ. California Press: Berkeley, Calif. 469-489 (1961b).

12. Parzen, E.: Nonparametric statistical data modeling. J. Amer. Statist. Assoc. 74, 105-120 (1979a).

13. Parzen, E.: A density-quantile function perspective on robust estimation. Robustness in Statistics edited by R. Lanner and G. Wilkinson, Academic Press: New York 237-258 (1979b).

14. Reid, N.: Influence functions for censored data. Ann. Statist. 9, 78-92 (1981).

15. Sacks, J. and Ylvisaker, D.: Designs for regression problems with correlated errors. Ann.Math. Statist. 37, 66-89 (1966).

16. Sander, J.: The weak convergence of quantiles of the product-limit estimator. Technical Report No. 5, Dept. Statist., Stanford Univ. (1975a).

17. Sander, J.: Asymptotic normality of linear combinations of functions of order statistics with censored data. Technical Report No. 8, Dept. Statist., Stanford Univ. (1975b).

18. Sander, J.: A comparison of three estimators derived from the Kaplan-Meier estimator. Technical Report No. 11, Dept. of Statist., Stanford Univ. (1975c).

19. Susarla, V. and Van Ryzin, J.: Large sample theory for an estimator of mean survival time from censored samples. Ann. Statist. 8, 1001-1016 (1980).

20. Weiss, L.: On estimating location and scale parameters from truncated samples. Naval Res. Logist. Quart. 11, 125-134 (1964).

21. Weiss, L. and Wolfowitz, J.: Asymptotically efficient non-parametric estimators of location and scale parameters. Z. Wahrscheinlichkeitstheorie verw. Geb. 16, 134-150. (1970).

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER 149	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Weighted L^2 Quantile Distance Estimators For Randomly Censored Data		5. TYPE OF REPORT & PERIOD COVERED TECHNICAL REPORT 6. PERFORMING ORG. REPORT NUMBER 149
7. AUTHOR(s) R. L. Eubank and V. N. LaRiccia		8. CONTRACT OR GRANT NUMBER(s) N00014-73-C-0439
9. PERFORMING ORGANIZATION NAME AND ADDRESS Southern Methodist University Dallas, Texas 75275		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Arlington, VA 22217		12. REPORT DATE November 1981 13. NUMBER OF PAGES 20
14. MONITORING AGENCY NAME & ADDRESS(If different from Controlling Office)		15. SECURITY CLASS. (of this report) 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

This document has been approved for public release and sale; its distribution is unlimited. Reproduction in whole or in part is permitted for any purposes of the United States Government.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Quantile function, minimum distance, estimation, random censoring, reproducing kernel Hilbert space, optimal weight function.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The asymptotic properties of a family of minimum quantile function distance estimators for randomly censored data sets are considered. These procedures produce an estimator of the parameter vector that minimizes a weighted L^2 distance measure between the Kaplan-Meier quantile function and an assumed parametric family of quantile functions. Regularity conditions are provided which insure that these estimators are consistent and asymptotically normal. An optimal weight function is derived for single parameter families, which, for location/scale families, results in censored sample analogs of estimators such as those suggested by Parzen (1979a, 1979b), and Weiss and Wolfowitz (1970).

DD FORM 1473 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)