A DENSITY-QUANTILE FUNCTION APPROACH TO ADAPTIVE LOCATION OR SCALE PARAMETER ESTIMATION

by

Randall L. Eubank

Technical Report No. 155
Department of Statistics ONR Contract

December, 1981

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

DEPARTMENT OF STATISTICS
Southern Methodist University
Dallas, Texas 757275

A DENSITY-QUANTILE FUNCTION APPROACH TO ADAPTIVE LOCATION OR SCALE PARAMETER ESTIMATION

Randall L. Eubank
Department of Statistics
Southern Methodist University
Dallas, Texas 75275

ABSTRACT: Two adaptive estimation procedures are developed for location or scale parameter estimation through the use of the asymptotically best linear unbiased estimator based on sample order statistics. The two procedures differ in their use of the data to guide optimal order statistic selection and are, consequently, termed partially and fully adaptive to indicate the associated degree of guidance. The partially adaptive procedure is developed in a data summary framework, similar to that utilized in exploratory data analysis, which provides a computationally simple scheme for obtaining estimators with high guaranteed asymptotic relative efficiency. The partially adaptive approach can also be viewed as providing techniques for adaptive data summary construction.

KEY WORDS

Adaptive estimation of location or scale parameter
Asymptotically best linear unbiased estimate
Sample quantiles
Asymptotically optimal spacings
Optimal spacing densities
Adaptive data summaries

1. INTRODUCTION

The location and scale parameter model assumes that a random sample, $X_1, \ldots X_n$, is obtained from a distribution function of the form $F(\frac{x-\mu}{\sigma})$ where μ and σ are respectively a location and scale parameter. Usually μ or σ requires estimation from the data. In this paper attention will be focused on the asymptotically best linear unbiased estimator (ABLUE) of μ or σ based on $k < n$ sample order statistics.

The objective of this paper is to construct an adaptive approach to location or scale parameter estimation which is based on the ABLUE and applicable to both symmetric and skewed distributions. Using techniques developed by Eubank (1981), two adaptive approaches are suggested. The first procedure, termed partially adaptive, entails the use of certain goodness of fit criteria and/or prior knowledge to identify a set of probability laws to which F is believed to belong. Then, a "best" set of order statistics is obtained for estimator construction. The entire procedure utilizes only one subsample from the entire set of sample quantiles and is developed in a data summary framework similar to that used in exploratory data analysis. The second procedure is termed fully adaptive as it requires the estimation, from the sample, of the density-quantile function (c.f. Parzen (1979)) prior to the selection of the order statistics used in estimator construction. Although both procedures are derived using large sample theory, work by Chan and Chan (1973) and Chan, Chan and Mead (1973 a,b) suggest that in certain cases these techniques may work well for even moderate or small n.

There are several reasons for using the ABLUE in the context of adaptive estimation. Perhaps the foremost of these, for the purposes of this paper, is that asymptotically (as $k \to \infty$) the location of optimal quantiles for the ABLUE can be completely charactized in a closed, easily used form for many symmetric and skewed distributions (c.f. Eubank (1981)). Such characterizations may be employed to obtain insight into the issues of robust and, as will be seen in Section 4, adaptive estimation. Other reasons for the use of ABLUE's are the high asymptotic relative efficiencies (ARE's) obtained from the ABLUE for even small k, when F is known, and the ease of estimator computation. In fact, many of the calculations in the partially adaptive procedure are readily accomplished without computer assistance.

Adaptive estimation of a location parameter by ABLUE's has been considered by Chan and Rhodin (1980). They assume that F is an element of the Tukey's lambda family or corresponds to the normal, double exponential, or Cauchy distribution. Thus, in contrast to this paper, their results apply only to symmetric distributions and location parameter estimation. Their approach is to first obtain three quantiles which are utilized to estimate a measure of tail length. The locations of the optimal quantiles for ABLUE construction are then found by maximizing the minimum ARE obtainable through an incorrect choice for F, over several distributions whose tail lengths lie within a neighborhood of this estimate. The ARE of the resulting estimator is called a guaranteed ARE (GARE). This approach necessitates the selection of two subsamples from the data at different times. In certain instances this could make their approach unfeasible (c.f. Eisenberger and Posner (1965)). Surveys and references on the general problem of adaptive estimation can be found in Hogg (1974) and Huber (1977).

The development of the adaptive procedures presented in Section 4 is facilitated by the use of results obtained by Eubank (1981). In Section 3 these results are reviewed and utilized to make certain distributional comparisons. Section 2 is devoted to background and notational preliminaries. Section 5 considers the implications of the contents of Sections 3 and 4 for robust location and scale parameter estimation.

2. THE DENSITY-QUANTILE FUNCTION AND THE ABLUE

In subsequent discussions it will be assumed that F admits a probability density function, f. The <u>quantile function</u> corresponding to f is $Q(u) = F^{-1}(u)$, $0 \leq u \leq 1$. The <u>density-quantile function</u> is defined as $d(u) = f(Q(u))$, $0 \leq u \leq 1$.

Using the sample order statistics, $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$, define the <u>sample quantile function</u> by

$$\tilde{Q}(u) = X_{(j)}, \quad \frac{j-1}{n} < u \leq \frac{j}{n}, \; j = 1,\ldots,n. \tag{2.1}$$

Given a spacing, $T = \{u_1,\ldots,u_k\}$, (k real numbers satisfying $0<u_1<u_2<\ldots<u_k < 1$) , the corresponding sample quantiles, $\tilde{Q}(u_1),\ldots,$ $\tilde{Q}(u_k)$, have been shown by Mosteller (1946) to have a normal limiting distribution. Ogawa (1951) has used the form of the limiting distribution to obtain formulae for the ABLUE's, $\mu^*(T)$ and $\sigma^*(T)$, of μ and σ, and their corresponding asymptotic relative efficiency with respect to the Cramér-Rao lower variance bound.

Let $T = \{u_1, \ldots, u_k\}$ be a spacing and let $u_0 = 0$, $u_{k+1} = 1$. Define the quantities K_1 and K_2 by

$$K_1\,(T;F) = \sum_{i=1}^{k+1} \frac{[d(u_i) - d(u_{i-1})]^2}{u_i - u_{i-1}} \tag{2.2}$$

and

$$K_2\,(T;F) = \sum_{i=1}^{k+1} \frac{[d(u_i)Q(u_i) - d(u_{i-1})Q(u_{i-1})]^2}{u_i - u_{i-1}} \quad . \tag{2.3}$$

The asymptotic relative efficiency formulas given by Ogawa are, in this notation,

$$ARE_F(\mu^*(T)) = \frac{K_1\,(T;F)}{E\left[\left(\frac{f'(X)}{f\,(X)}\right)^2\right]} \tag{2.4}$$

and

$$ARE_F(\sigma^*(T)) = \frac{K_2(T;F)}{E\left[\left(X\,\frac{f'(X)}{f\,(X)}\right)^2\right] - 1} \tag{2.5}$$

where ARE_F denotes ARE under the assumption that F is the true distribution function.

To obtain optimal estimators, T should be chosen to maximize one of (2.4) or (2.5). A spacing which results in a maximum for one of these expressions is termed an <u>optimal spacing</u>.

Chan and Rhodin (1980) have obtained spacings which are robust in the sense that they provide high ARE's over several, user specified, distributional forms. In contrast, the objective of this paper is to provide an adaptive approach to spacing selection. Consequently, the procedures which will be developed will utilize the data, either fully or in part, to guide the spacing selection process.

In the next section a spacing selection procedure will be discussed which, when F is known, can be utilized to obtain spacings which correspond to near maximum ARE's. It will then be seen that these spacings exhibit similiar behaviour for certain distribution types. This latter fact provides the basis for the adaptive techniques presented in Section 4.

3. ASYMPTOTICALLY OPTIMAL SPACINGS

Parzen (1979) has shown that location and scale parameter estimation by linear functions of order statistics can be considered as a regression analysis problem for continuous parameter time series through use of the model

$$d(u)\tilde{Q}(u) = \mu d(u) + \sigma d(u)Q(u) + \sigma_B B(u)\,, \quad u \in [0,1]\,, \tag{3.1}$$

where $\sigma_B = \sigma/\sqrt{n}$ and $B(\cdot)$ is a Brownian bridge process. If samples are taken from this model at a finite number of design points, μ and σ may be estimated through the use of generalized least squares. Eubank (1981) has shown that the resulting estimators and corresponding variance formulae agree with those for the ABLUE given by Ogawa (1951). Therefore, the problem of optimal regression design, in the minimum variance sense, for model (3.1) is identical to the optimal spacing problem.

Using asymptotic theory for regression designs for continuous parameter time series, Eubank (1981) has shown that certain density functions on [0,1] can be used to obtain spacings which have "good properties" which will be discussed subsequently. These densities are

$$h(u) = \begin{cases} \dfrac{(d(u)")^{2/3}}{\int_0^1 (d(s)")^{2/3} ds} & \text{, when } \sigma \text{ is unknown,} \\[2ex] \dfrac{([d(u)Q(u)]")^{2/3}}{\int_0^1 ([d(s)Q(s)]")^{2/3} ds} & \text{, when } \mu \text{ is known.} \end{cases} \qquad (3.2)$$

Let H denote the distribution function corresponding to h. Then the k-element spacing generated by H is $T_k = \{H^{-1}\left(\frac{1}{k+1}\right), H^{-1}\left(\frac{2}{k+1}\right), \ldots, H^{-1}\left(\frac{k}{k+1}\right)\}$.

To be more precise about the "good properties" these spacings possess, consider the problem of location parameter estimation when σ is known. If h denotes the appropriate density in (3.2), then, by successively increasing k, h may be used to generate a sequence of spacings, $\{T_k\}_{k=1}^{\infty}$. This sequence satisfies

$$\lim_{k \to \infty} \frac{1 - \text{ARE}\ (\mu^*(T_k))}{1 - \inf_{T \in D_k} \text{ARE}(\mu^*(T))} = 1 \qquad (3.3)$$

where D_k denotes the set of all k element spacings. For scale parameter estimation an analogue of (3.3) holds with μ^* replaced by σ^*.

Equation (3.3) has the important implication that the spacings generated by h will, for large k, behave like the optimal spacings with respect to ARE's. The spacings in a sequence which satisfy (3.3) are termed asymptotically optimal spacings. The densities in (3.2) are termed optimal spacing densities.

Since (3.3) is only an asymptotic result, caution must be utilized when discussing the properties of asymptotically optimal spacings for finite k. However, for most distributional forms, the asymptotically optimal spacings have ARE's quite close to those of their optimal counterparts even for k as small as 3. This remark is readily verified through reference to Table 1 where the difference in efficiencies between optimal and asymptotically optimal spacings are given for various choices of F when k = 3, 7.

In addition to being used for actual spacing computation, the h (or equivalently H or H^{-1}) functions are useful tools for distributional comparisons since they provide (asymptotically) a characterization of the behaviour of the optimal spacings. Since the H^{-1} function is required for spacing computation it will be more convenient to compare the H^{-1}, rather than h, functions for various choices of F.

The H^{-1} function dictates how the elements of asymptotically optimal spacings are concentrated for a particular distribution type. Figures 1 and 2 show H^{-1} functions for several common distribution types for the case of either σ or μ known. Through examination of these figures it is possible to recognize three common types of behaviour for H^{-1}:

(1) Uniform: $H^{-1}(u) = u$ (e.g. the logistic when σ is known and the Pareto with μ known and shape parameter, ν, equal to 1).

TABLE 1. Loss in ARE from the use asymptotically optimal rather than optimal spacings

Distribution	Parameter to be estimated	Loss in ARE[a] k = 3	Loss in ARE[a] k = 7
Normal	μ	.0092	.0017
	σ	.0107	.0182
Logistic	μ	0[b]	0[b]
	σ	.0152	---[c]
Cauchy	μ	.0581	.0116
	σ	.0524	.0078
Extreme Value	μ	.0143	.0013
Exponential	σ	.0143	.0013
Pareto $\nu = .5$	σ	.0006	.0001
$\nu = 1$	σ	.001	.0001
$\nu = 3$	σ	.0025	.0002
Lognormal	σ	.0092	.0017

(a) References and values for ARE's pertaining to the optimal spacings as well as the ARE's for asymptotically optimal spacings may be found in Eubank (1979).

(b) The asymptotically optimal spacings coincide with the optimal spacings in this instance.

(c) The optimal spacings for this case are, to the author's knowledge, unknown.

Figure 1. The H^{-1} functions for various distributions in the case that σ is known.

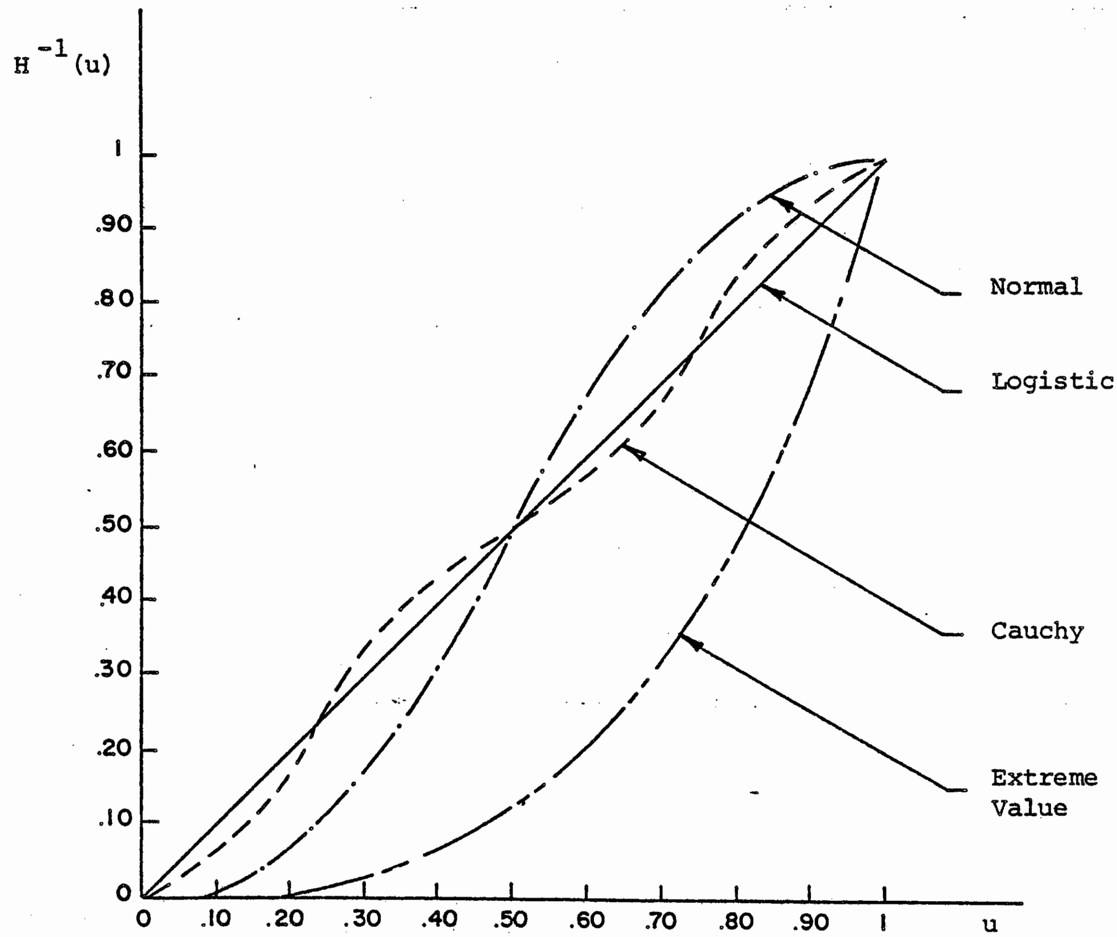
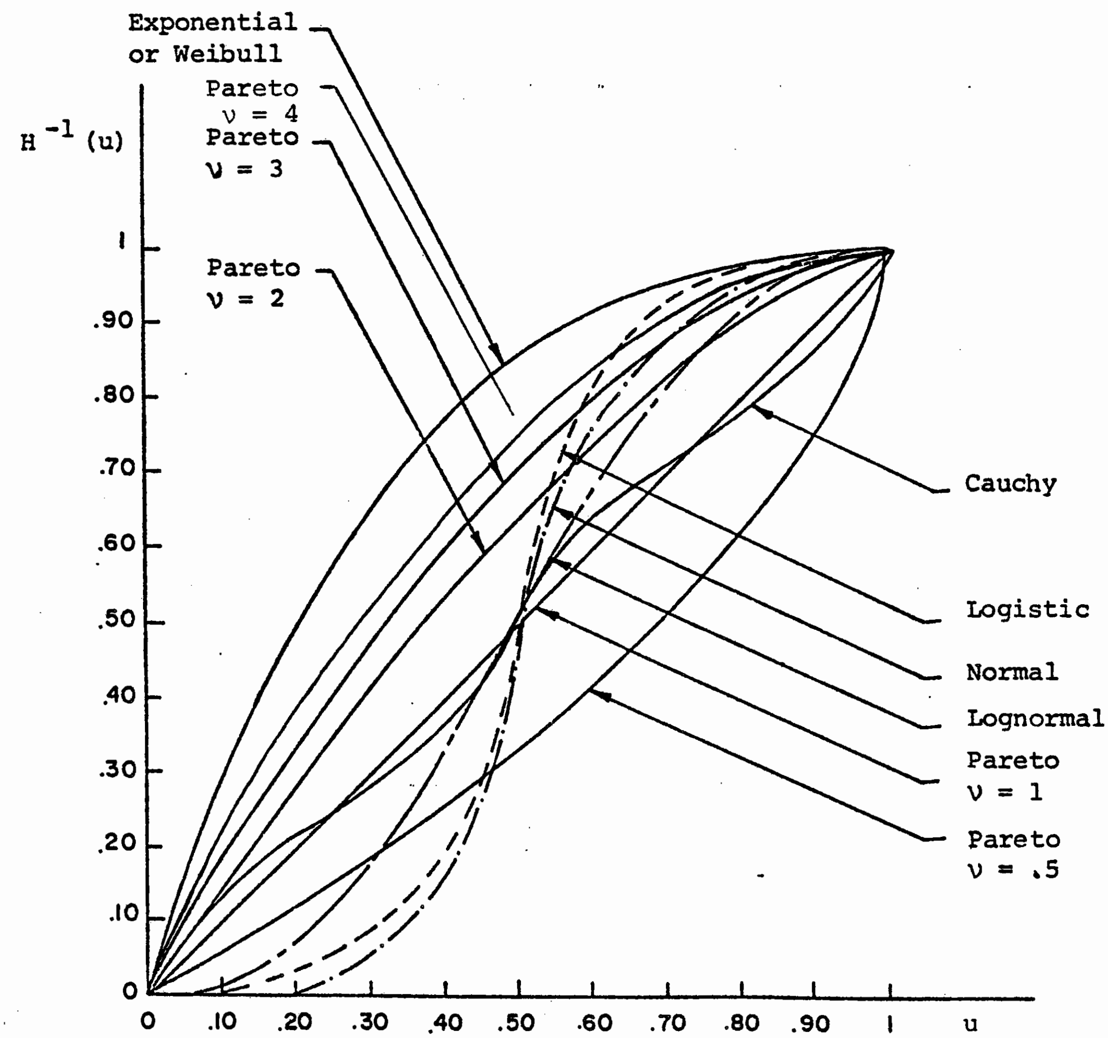


Figure 2 - The H^{-1} functions for various distributions in the case that μ is known.



(2) Skewed: In this case H^{-1} often behaves like u^3 or $1-(1-u)^3$ (e.g., the extreme value when σ is known and the exponential when μ is known).

(3) Symmetric: $1-H^{-1}(u) = H^{-1}(1-u)$ (e.g. the normal or logistic when μ or σ is known).

Since H^{-1} is the quantile function corresponding to H its characteristics may be interpreted as one would interpret those of any quantile function (c.f. Parzen (1979)). Thus many spacing elements will be selected from an interval (on the H^{-1} axis) which corresponds to a relatively flat shape (near 0 slope) for H^{-1} while intervals which correspond to sharp rises will have low spacing element concentration. For example, the spacing elements for the exponential, when μ is known, will be concentrated near 1 with few elements near 0. This will entail selection from among the larger, rather than the smaller, order statistics for the estimation of σ from exponential data. When H^{-1} is symmetric spacing elements are equally and symmetrically distributed about .5.

In the process of constructing a partially adaptive approach to estimation based on the ABLUE (see Section 4) it becomes necessary to compare the various H^{-1} functions. Through grouping together distributions whose H^{-1} functions have similar shapes it is found that by slightly altering the asymptotically optimal spacings it is possible to suggest spacings which work well for several distributions simultaneously. These perturbed or altered spacings are nearly as efficient as the asymptotically optimal or optimal ones due, in part, to the continuity of $K_i(T;F)$, $i = 1,2$, as a function of T (continuity is readily established by reference to the work of Sacks and Ylvisaker (1966)). Thus, if $T^1 = (t_1^1,\ldots,t_k^1)$ and $T^2 = (t_1^2,\ldots,t_k^2)$ are two spacings, the difference $|K_i(T^1;F) - K_i(T^2;F)|$ will be small provided $\max_{1\leq j\leq k} |t_j^1 - t_j^2|$ is also sufficiently small. Precise bounds for this difference, or equivalently for the loss in ARE, may be obtained by imposing smoothness conditions on K_i or by examining specific classes of F's.

4. ADAPTIVE LOCATION OR SCALE PARAMETER ESTIMATION

In this section adaptive data analysis procedures are presented which are based on ABLUE's. Both fully and partially adaptive location or scale parameter estimation schemes are considered and, as a consequence of the development of the partially adaptive approach, techniques for adaptive data summary construction are discussed.

Since the properties of the ABLUE derive from the limiting distribution for the sample quantiles, the results presented in this section are designed for use, primarily with large data sets. It is therefore important to note that the subsequent adaptive techniques deal simultaneously with the issues of estimation and data compression.

Adaptive data summaries and a partially adaptive estimation procedure.

Data summaries consisting of 7, or fewer, sample order statistics are frequently utilized in exploratory data analysis (c.f. Parzen (1979)). The construction of a data summary may be an end in itself or, only a preliminary step before transformation or re-expression of the entire data set. In either instance the estimation of location or scale parameters, from the quantiles which compose the summary, plays an important role in the analysis of the data. Thus, it is reasonable to consider the construction of data summaries which, among other things, are capable of providing good estimates of μ and σ. The partially adaptive estimation procedure presented below provides such summaries.

Using the results of Section 3 it is possible to develop an adaptive approach to data summary construction. The quantiles indicated in Tables 2 provide what is termed a <u>19 number adaptive data summary</u>. The values indicated in Tables 2 contain as subsets various 7 number sub-summaries which may be utilized to summarize data from each of the distribution considered in this paper. The sub-summaries appropriate for any particular distribution are composed of quantiles which may be utilized to obtain good estimates of either μ or σ. For this reason, Parzen (1980) has termed the quantiles in Table 2 a "universal data summary".

To use Table 2 to construct a data summary the following procedure may be utilized. First, the 19 quantiles, consisting of the median, the j/16 percentiles, j = 1,...,7, and the .01 and .02 percentiles are obtained from the data. These quantiles are then used to determine one or more laws among those listed in Tables 3 or 4, which seem to fit the data through the use, perhaps, of goodness-of-fit procedures. The goodness-of-fit techniques developed by Parzen (1979) are perhaps best suited to this purpose since their use does not entail the estimation of μ or σ. Let D denote this set or laws. The ARE's in Table 3 or 4 can then be used to find which element of D maximizes

$$\min_{G \in D} \text{ARE}_F(\cdot|G) \, , \tag{4.1}$$

where $\text{ARE}_F(\cdot|G)$ denotes the ARE obtained by using the spacing corresponding to F when G is the true d.f. and the "$\cdot$" is used as a place

TABLE 2. An Adaptive 19 Number Data Summary

1. The median, $\tilde{Q}(.5)$.
2. The $\frac{7}{16}$ percentile, $\tilde{Q}(.4375)$ and $\tilde{Q}(.5625)$.
3. The $\frac{3}{8}$ percentile, $\tilde{Q}(.375)$ and $\tilde{Q}(.625)$.
4. The $\frac{5}{16}$ percentiles, $\tilde{Q}(.3125)$ and $\tilde{Q}(.6875)$.
5. The quartiles, $\tilde{Q}(.25)$ and $\tilde{Q}(.75)$.
6. The $\frac{3}{16}$ percentiles, $\tilde{Q}(.1875)$ and $\tilde{Q}(.8125)$.
7. The $\frac{1}{8}$ percentiles, $\tilde{Q}(.125)$ and $\tilde{Q}(.875)$.
8. The $\frac{1}{16}$ percentiles, $\tilde{Q}(.0625)$ and $\tilde{Q}(.9375)$.
9. The $\frac{2}{100}$ percentiles, $\tilde{Q}(.02)$ and $\tilde{Q}(.98)$.
10. The $\frac{1}{100}$ percentiles, $\tilde{Q}(.01)$ and $\tilde{Q}(.99)$.

TABLE 3. The $ARE_F(\mu^*(T)|G)$ for various combinations of F and G

F/G	Normal	Cauchy	Logistic	Extreme Value
Normal	.9622	.9264	.9713	.9364
Cauchy	.9450	.9496	.9844	.8663
Logistic	.9450	.9469	.9844	.8663
Extreme Value	.8978	.8464	.9450	.9653

Table 4. The $ARE_F(\sigma^*(T)|G)$ for various combinations of F and G.

F \ G	Exponential or Weibull	Pareto ν=.5	Pareto ν=1	Pareto ν=2	Pareto ν=3	Pareto ν=4	Logistic	Normal	Lognormal	Cauchy
Exponential or Weibull	.9653	.8965	.9450	.9627	.9627	.9687	.6471	.5968	.8978	.8261
Pareto ν = .5	.6849	.9838	.9609	.8966	.8514	.8211	.6027	.5213	.8985	.8600
Pareto ν = 1	.8863	.9805	.9844	.8723	.9574	.9449	.7510	.6394	.9450	.9496
Pareto ν = 2	.9251	.9644	.9785	.9792	.9743	.9693	.7461	.6479	.9401	.9374
Pareto ν = 3	.9563	.9583	.9743	.9770	.9757	.9741	.7688	.6958	.9415	.9204
Pareto ν = 4	.9611	.9380	.9670	.9757	.9767	.9761	.7090	.6393	.9244	.9001
Logistic	.9354	.9306	.9348	.9391	.9407	.9411	.9009	.8926	.9453	.7516
Normal	.9001	.8825	.8922	.8959	.8899	.8994	.8393	.8538	.9078	.5413
Lognormal	.9364	.9671	.9713	.8673	.9631	.9597	.8834	.8555	.9622	.8928
Cauchy	.8061	.9655	.9727	.9508	.9270	.9085	.6230	.5006	.9120	.9413

holder for the appropriate one of $\mu^*(T)$ or $\sigma^*(T)$ depending on whether the objective is location or scale parameter estimation. If F^* is the distribution which maximizes (4.1) then the 7 element data summary is obtained by selecting those of the initial 19 quantiles which correspond to the check marked spacings under the F^* column in the apporpriate one of Tables 5 or 6.

The partially adaptive estimation procedure consists of using the quantiles obtained for the 7 number data summary to construct an estimator of μ or σ. This is readily accomplished through using the coefficients for the quantiles presented in Tables 7. Table 7 also gives the ARE for the estimator assuming F^* is the true distribution. In contrast, the GARE for the estimator is the minimal value of $ARE_{F^*}(\cdot|G)$ over $G \varepsilon D$. This value was found in determining the appropriate subsummary.

EXAMPLE. To illustrate the use of the partially adaptive procedure two numerical examples will be considered. In each example, the approach proposed by Parzen (1979) for ascertaining the goodness of fit of a probability law is utilized to determine a set of likely candidates, D, for the parent distribution of the data. The necessary computations were conducted using the ONESAM statistical package developed by Parzen and Anderson (1980).

The elements of the Parzen approach which are important for this discussion may be summarized as follows. Let F_0 be a hypothesized distribution with corresponding density-quantile function d_0. Under the null hypothesis, that F_0 is the parent distribution of the sample, the function

TABLE 5. Order Statistic Selection for Location Parameter Estimation by Seven Order Statistics

	Distribution			
Spacing	Normal	Cauchy	Logistic	Extreme Value
.01				√
.02	√			√
.0625				√
.125	√	√	√	√
.1875				
.25		√	√	√
.3125	√			
.375		√	√	
.4375				√
.5	√	√	√	
.5625				
.625		√	√	
.6875	√			√
.75		√	√	
.8125				
.875	√	√	√	
.9375				
.98	√			
.99				

Table 6. Order Statistic Selection for Scale Parameter Estimation by Seven Order Statistics

Spacing	Exponential or Weibull	Pareto ν =.5	Pareto ν= 1	Pareto ν = 2	Pareto ν = 3	Pareto ν = 4	Logistic	Normal	Lognormal	Cauchy
					DISTRIBUTION					
.01							√	√		
.02								√	√	
.0625		√					√			
.125			√					√	√	
.1875		√		√	√		√			√
.25		√	√			√				√
.3125	√	√		√					√	√
.375			√		√					
.4375		√				√				
.5			√	√			√	√	√	√
.5625	√	√			√					
.625			√	√		√				
.6875		√			√				√	√
.75	√		√	√		√				√
.8125					√		√			√
.875	√		√	√		√	√	√		
.9375	√			√	√	√	√			
.98	√				√	√		√	√	
.99	√						√	√		

Table 7. Coefficients and Efficiencies for the ABLUE based on Seven Order Statistics

b_1	b_2	b_3	b_4	b_5	b_6	b_7	b_0	ARE
			a. Normal Distribution, σ Known					
.0464	.1517	.2026	.1986	.2026	.1517	.0464	.0	.9622
			b. Normal Distribution, μ Known					
-.0383	-.0691	-.2338	.0	.2338	.0691	.0383	.0	.8538
			c. Exponential Distribution, μ Known					
.2906	.2252	.1587	.0898	.055	.0204	.0144	.854	.9653
			d. Pareto Distribution, ν = .5, μ Known					
.5398	.3195	.1508	.1548	.0954	.0349	.0127	1.3079	.9838
			e. Pareto Distribution, ν = 1, μ Known					
.5833	.4286	.2976	.1905	.1071	.0477	.0119	1.6667	.9844
			f. Pareto Distribution, ν = 2, μ Known					
.7698	.6708	.4717	.2893	.1947	.0684	.0264	2.4912	.9792
			g. Pareto Distribution, ν = 2, μ Known					
1.211	.9448	.6031	.3995	.2934	.0779	.0154	3.4549	.9757
			h. Pareto Distribution, ν = 4, σ Known					
1.4387	1.0941	.7184	.4946	.2383	.1006	.0337	4.1184	.9761
			i. Cauchy Distribution, σ Known					
-.0518	.0	.3018	.5	.3018	.0	.0518	.0	.9496
			j. Cauchy Distribution, μ Known					
-.1232	-.1311	-.276	.0	.276	.1311	.12	.0	.9463
			k. Logistic Distribution, σ Known					
.0833	.1429	.1786	.1905	.1786	.1429	.0833	.0	.9844
			l. Logistic Distribution, μ Known					
-.0334	-.1351	-.2644	.0	.2644	.1351	.0334	.0	.9009
			m. Weibull Distribution, γ = 4/3, μ Known					
.2274	.2147	.1722	.1078	.071	.0286	.021	.8427	.9653
			n. Weibull Distribution, γ = 2, μ Known					
.1779	.2047	.1868	.1294	.0916	.0403	.0308	.8616	.9653
			o. Weibull Distribution, γ = 4, μ Known					
.1392	.1952	.2027	.1554	.1183	.0566	.0451	.9126	.9653
			p. Extreme Value Distribution, σ Known					
.0661	.0796	.1526	.1866	.22	.1861	.1089	-.4314	.9653
			q. Lognormal Distribution, μ Known					
.3615	.4793	.3303	.1986	.1243	.048	.0059	1.548	.9622

$r(t) = d_0(t)/(d(t)\{\int_0^1 \frac{d_0(x)}{d(x)} dx\})$ will be identically one over all of [0,1]. To utilize Parzen's procedure one first forms, from the data, raw estimators of $\tilde{r}$ and its corresponding distribution function, $\tilde{R}$, which will be denoted by $\tilde{r}$ and $\tilde{R}$ respectively. Then a variety of diagnostic statistics based on $\tilde{r}$ and $\tilde{R}$ are available for judging the goodness of Fit of F_0. In particular, one may consider how well $\tilde{R}$ conforms to the uniform distribution of [0,1]. Further anaysis is provided by determination of the optimal order, m, of an autoregressive smoother of $\tilde{\tilde{r}}$. One approach to the selection of m is through the use of the CAT criterion. When CAT chooses m = 0 this is regarded as additional confirmation that H_0 holds.

As a first example consider estimating a location parameter from the data on lifetimes, in hours, of 417 40-watt internally frosted incandescent lamps studied by Davis (1952). The 19 sample quantiles listed in Table 2 were obtained and analyzed using the Parzen procedure. Although CAT selected order zero for all four of the laws, through comparison of $\tilde{\tilde{R}}$ to the uniform distribution it is found that the normal and logistic distributions provide far superior fits. Therefore, an appropriate choice would be D = {normal, logistic}.

From Table 3 it is seen that when F is the normal distribution

$$ARE_F(\mu^*(T)\,|\,\text{Normal}) = .9622$$

$$ARE_F(\mu^*(T)\,|\,\text{Logistic}) = .9713$$

and that when F is the logistic distribution

$$ARE_F(\mu^*(T)\,|\,\text{Normal}) = .9450$$

$$ARE_F(\mu^*(T)\,|\,\text{Logistic}) = .9844 .$$

As the choice of the normal distribution results in the largest minimum ARE over D, the estimate of μ is constructed using the spacings in Table

5 for the normal distribution and their corresponding coefficients in Table 7. The resulting estimator is

$$
\begin{aligned}
\mu^*(T) &= .0464\tilde{Q}(.02) + .1517\tilde{Q}(.125) + .2026\tilde{Q}(.3125) \\
&\quad + .1986\tilde{Q}(.5) + .2026\tilde{Q}(.6875) + .1517\tilde{Q}(.875) \\
&\quad + .0464\tilde{Q}(.98) \\
&= .0464(623) + .1517(836) + .2026(954) \\
&\quad + .1986(1037) + .2026(1127) + .1517(1243) \\
&\quad + .0464(1461) \\
&= 1039.6407
\end{aligned}
$$

which is close to the value 1044.3 given by Chan and Rhodin (1980) as their estimate of μ for this data.

A test for normality performed by Davis on all 417 data values found the normal to be acceptable although a test for kurtosis indicated that the observations were more peaked at the mean and flatter in the tails. Indeed, if the entire set of values is analyzed using the Parzen procedure the normality hypothesis is rejected and the logistic is seen to be the best choice for the distribution of this data from among those laws given in Table 3. The logistic was also used by Chan and Rhodin in obtaining a robust estimator of μ. As the partially adaptive technique uses only 19 sample quantiles it is not surprizing that both the normal and logistic were selected as the members of D. It is encouraging, however, that the choices indicated for D included the "correct" model and that the resulting estimator of μ is close to the value given by other authors using the entire sample for their analysis.

Now consider the estimation of a scale parameter from the data on 109 time intervals, in days, between explosions in mines involving more than ten men killed from December 6, 1875 to May 29, 1951, given by Maguire, Pearson and Wynn (1952). To determine the membership of D the 19 sample quantiles in Table 2 were used to ascertain the goodness of fit of several distribution candidates, including the exponential, several Weibull distributions with shape parameters close to the exponential and the Pareto with shape parameters 2, 3 and 4. Examination of $\tilde{R}$ indicated that the data was well modeled by either of the three Pareto distributions which were considered. In addition these were the only instances when CAT selected an optimal order of zero. Hence, in this case, D was taken to consist of three members, namely the Pareto with shape parameters ν = 2, 3 and 4.

From Table 4 it is seen that, for the Pareto

$$ARE_{\nu = 2}(\sigma^*(T)|\nu = 2) = .9792$$

$$ARE_{\nu = 2}(\sigma^*(T)|\nu = 3) = .9743$$

$$ARE_{\nu = 2}(\sigma^*(T)|\nu = 4) = .9693$$

so that

$$\min_{G \epsilon D} ARE_{\nu=2}(\sigma^*(T)|G) = .9693$$

Similarly for ν = 3 and ν = 4 the values for $\min_{G \epsilon D} ARE_{\nu}(\sigma^*(T)|G)$ are .9741 and .9757 respectively. As the Pareto with ν = 4 has the largest minimum ARE over the laws in D its corresponding spacings and coefficients in Tables 6 and 7 are used in estimating σ. The resulting estimator is

$$\begin{aligned}\sigma^*(T) &= 4.1184 + 1.4387\tilde{Q}(.25) + 1.0941\tilde{Q}(.4375) \\ &\quad + .7184\tilde{Q}(.625) + .4946\tilde{Q}(.75) + .2383\tilde{Q}(.875) \\ &\quad + .1006\tilde{Q}(.9375) + .0337\tilde{Q}(.98) \\ &= 4.1184 + 1.4387(54) + 1.0941(120) + .7184(217) \\ &\quad + .4946(336) + .2383(390) + .1006(745) \\ &\quad + .0337(1357) \\ &= 748.7935 \quad .\end{aligned}$$

If the entire data set is analyzed using the Parzen procedure the Pareto with $\nu = 2$ is eliminated from D. The end result, in terms of the value of the estimator, will of course remain the same as the Pareto with $\nu = 4$ still has the largest minimum ARE. It is noteworthy that the set D identified using the 19 quantiles again includes the laws which were indicated as appropriate for this data through use of the entire sample.

Some remarks regarding the procedure utilized in selecting the quantiles presented in Table 2 now seem appropriate. The quantiles suggested for the 19 number adaptive data summary were selected through careful inspection of the sets of 7 element asymptotically optimal spacings for the various distributions of interest. In many instances, as suggested by Figures 1 and 2, these sets had many elements which were quite similar. By identifying these similar elements and then slightly altering them to obtain a common value, an overall reduction in the number of spacings, and hence the number of quantiles, required to accurately summarize many distributions simultaneously was achieved. In determining the common value to use for several similar spacing elements, preference was given, when applicable to values frequently suggested for data summaries such as the quartiles and eighths. The perturbation of the spacing elements was expected to have minimal detrimental impact on the ARE's of the corresponding estimators due to the continuity of these functions as discussed in Section 3. That this expectation was realized can be seen from examination of the ARE's presented in Table 7. Due to the nature of their construction, it is not possible to attach any overall optimality properties to the adaptive data summary given in Table 2. However, the suggested summary is appealing since it works well and tends to agree with the values frequently suggested in data summary construction.

Fully adaptive approach. The fully adaptive estimation procedure requires the use of an estimator for the density-quantile function. Given $m \leq n$ sample quantiles such an estimator may be obtained through the use of the approach developed by Parzen (1979).

The fully adaptive location or scale parameter estimation procedure is summarized as follows:

1. Construct as estimator, $\hat{d}$, of the density-quantile function, d, based on $m \leq n$ of the sample quantiles.
2. Form the appropriate density in (3.2) using the estimator $\hat{d}$. To estimate $Q \cdot d$ use the relationship $Q' = 1/d$.
3. Let $\hat{h}$ denote the density constructed in Step 2 and let $\hat{H}$ represent its associated distribution function. To obtain an estimator of μ or σ apply the formulas given by Ogawa (1951) to the sample quantiles that correspond to the spacing $\{\hat{H}^{-1}(\frac{1}{k+1}),\ldots,\hat{H}^{-1}(\frac{k}{k+1})\}$ (c.f. also Eubank (1979) for the Ogawa formulas written in density-quantile function notation).

At present there seems to be no rule of thumb to use in the selection of m or the spacing for the initial m sample quantiles. However, if a reasonable guess is available for the form of F, one might select the m quantiles through use of the corresponding H^{-1} function. In such instances $m = k$ is a good choice, since, should it be found that $\hat{H}^{-1}$ is similar to H^{-1}, some or, perhaps, all further quantile selection could be avoided.

5. APPLICATION TO ROBUST ESTIMATION

Chan and Rhodin (1980) have considered the problem of robust location parameter estimation using ABLUE's. They consider estimators which are robust relative to a set of laws D, where robustness is measured by the estimators GARE. As a rule one would expect GARE to be adversely influenced by diversity in the members of D. This point is readily illustrated in the situation they consider, namely, when D is composed of members of the Tukey lambda family with different shape parameters. For instance, if D consists of distributions whose shape parameters are -.2, -.1 and 0 the GARE for location parameter estimation is found, from Table 3 of their paper to be .8889 whereas if the shape parameters are -.1, 0 and .1 the GARE is .8378.

A natural way to determine the influence of the diversity of D on the GARE is to compare the H^{-1} functions for the various members of D. A marked dissimilarity among H^{-1} functions would indicate that a high GARE is likely not obtainable since (asymptotically) the location of the optimal spacings would be quite different. Thus, ABLUE's should only be expected to be robust with respect to D if the elements of D have similar H^{-1} functions. The comparison of H^{-1} functions should, therefore, prove useful in determining reasonable choices for D.

6. SUMMARY AND CONCLUSIONS

The objective of this paper is to develop adaptive estimation procedures for location and scale parameters which are based on linear functions of order statistics. This is accomplished through the use of the ABLUE and the H^{-1} function which asymptotically characterizes the distribution of the elements of the optimal spacings for the sample quantiles. The applications of the H^{-1} function to robust estimation are also considered.

The two adaptive estimation procedures discussed in this paper differ in their dependence on and use of the data. The fully adaptive procedure entails the actual estimation of H^{-1} and requires the user to sample twice from the original data set. In contrast, use of the partially adaptive procedure requires only that one initial 19 quantiles subsample be obtained. Thus, in instances when repeated use of the data is not possible, the partially adaptive procedure may be preferable.

The focus of this paper has been entirely on single parameter estimation situations for certain distribution types. This approach can, however, be extended to include simultaneous parameter estimation in many instances and other distribution types such as the Gamma, and Tukey's lambda.

The concept of adaptive data summaries was developed as a consequence of the partially adaptive estimation procedure. For large data sets such summaries should provide a useful means of data compression. The advantage of adaptive data summaries in this context is that the data anlyst can obtain a summary which "fits" the data without the need for rigorous analysis and modeling of the entire data set.

REFERENCES

Chan, L. K. and Chan, N.N. (1973). On the optimum best linear unbiased estimates of the parameters of the normal distribution based on selected order statistics. Skand. Aktuarietidskr., 56, 120-8.

Chan, L. K. and Chan, N.N. and Mead, E. R. (1973a). Best linear unbiased estimates of the parameters of the logistic distribution based on selected order statistics. J. Amer. Statist. Assoc., 66, 889-92.

Chan, L. K., Chan, N.N. and Mead, E. R. (1973b). Linear estimation of the parameters of the Cauchy distribution using selected order statistics. Utilitas Math., 3, 311-318.

Chan, L.K. and Rhodin, L.S. (1980). Robust estimation of location using optimally chosen sample quantiles. Technometrics, 22, 225-37.

Davis, D. J. (1952). An analysis of some failure data. J. Amer. Statist. Assoc., 47, 113-50.

Eisenberger, I. and Posner, E. C. (1965). Systematic statistics used for data compression in space telemetry. J. Amer. Statist. Assoc., 60, 97-133.

Eubank, R. L. (1979). A density-quantile function approach to choosing order statistics for the estimation of location and scale parameters. Technical Report No. A-10, Institute of Statistics, Texas A & M University.

Eubank, R. L. (1981). A density-quantile function approach to optimal spacing selection. Ann. Statist., to appear.

Hogg, R. V. (1974). Adaptive robust procedures: a partial review and some suggestions for future applications and theory (with comments). J. Amer. Statist. Assoc., 69, 909-27.

Huber, P. J. (1977). Robust Statistical Procedures. SIAM Regional Conference Series in Applied Mathematics. Bristol, England: J. W. Arrowsmith Ltd.

Maguire, B. A., Pearson, E. S., and Wynn, A.H.A. (1952). The time interval between industrial accidents. Biometrika, 39, 168-180.

Mosteller, F. (1946). On some useful inefficient statistics. Ann. Math. Statist., 17, 175-213.

Ogawa, J. (1951). Contributions to the theory of systematic statistics, I., Osaka Math. J., 3, 131-142.

Parzen, E. (1979). Nonparametric statistical data modeling. J. Amer. Statist. Assoc., 74, 105-21.

Parzen, E. (1980). Data modeling using quantile and density-quantile functions. To appear in the *Proceeding of the Symposium on Recent Advances in Statistics* sponsored by the Lisbon Academy of Sciences, April, 1980, Lisbon, Portugal.

Parzen, E. and Anderson, S. (1980). Onesam, a computer program for non-parametric data analysis and density-quantile estimation. Technical report No. B-1, Institute of Statistics, Texas A & M University.

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

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 155	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) A Density-Quantile Function Approach to Adaptive Location or scale parameter Estimation		5. TYPE OF REPORT & PERIOD COVERED TECHNICAL REPORT 6. PERFORMING ORG. REPORT NUMBER 155
7. AUTHOR(s) Randall L. Eubank		8. CONTRACT OR GRANT NUMBER(s) N00014-75-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 December 1981 13. NUMBER OF PAGES 30
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)

Adaptive estimation of location or scale parameter; asymptotically best linear unbiased estimate; sample quantiles; asymptotically optimal spacings; optimal spacing densities; adaptive data summaries.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number) Two adaptive estimation procedures are developed for location or scale parameter estimation through the use of the asymptotically best linear unbiased estimator based on sample order statistics. The two procedures differ in their use of the data to guide optimal order statistic selection and are, consequently, termed partially and fully adaptive procedure is developed in a data summary framework, similar to that utilized in exploratory data anlysis, which provided a computationally simple scheme for obtaining estimators with high guaranteed asymptotic relative efficiency. The partially adaptive approach can also be viewed as providing techniques for adaptive data summary construction.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)