Statistische Hefte 25, 69–82 (1984)

Statistische Hefte
Statistical Papers

A family of minimum quantile distance estimators for the three-parameter Weibull distribution

T. J. Carmody, R. L. Eubank and V. N. LaRiccia

A family of minimum quantile distance estimators, based on a subset of the sample quantiles, is proposed for the parameters of the three-parameter Weibull distribution. The estimation procedure is applicable to either complete or censored samples and, through use of the associated distance measure, provides a goodness-of-fit test for the Weibull model. The proposed estimators are both consistent and asymptotically normal and, in a particular instance, are optimal over the class of all estimators based on the same quantile subset. The problem of optimal quantile selection is also considered.

1. Introduction

Let $X_{(1)},\ldots,X_{(n)}$ denote an ordered random sample from the three-parameter Weibull distribution having cumulative distribution function (c.d.f.)

$$G(x;\underline{\theta}) = 1 - \left\{\exp -\left(\frac{x-\mu}{\sigma}\right)^c\right\}, \quad x > \mu, \qquad (1.1)$$

where $\underline{\theta} = (\mu,\sigma,c)'$ with σ, $c > 0$ and $-\infty < \mu < \infty$. In this paper we consider the problem of simultaneous estimation of the elements of $\underline{\theta}$.

The three-parameter Weibull distribution has occupied an important role in areas such as reliability, biological modeling and queueing theory. Consequently, there is an extensive literature on the estimation of its parameters. Many of the important references are provided in Johnson and Kotz (1970), Kübler (1979) and Mann, Schafer and Singpurwalla (1974). In contrast to previous work, however, we propose the use of a minimum quantile distance estimator of $\underline{\theta}$. Minimum distance estimators have been extensively studied in recent years and have been found to possess commendable efficiency and robustness properties in many instances. The reader is referred to Parr and Schucany (1980) for an excellent discussion of minimum distance estimation using the empirical c.d.f. and to LaRiccia (1982) for a quantile domain approach.

The estimators of $\underline{\theta}$ considered in this paper can be described as follows. Let

$$Q(u) = -\ln(1-u) \ , \ 0 < u < 1, \tag{1.2}$$

and observe that the quantile function corresponding to $G(x;\underline{\theta})$ is

$$Q(u;\underline{\theta}) \equiv \inf_x\{x:G(x;\underline{\theta}) \geq u\} = \mu + \sigma Q(u)^{1/c}. \tag{1.3}$$

Define the sample quantile function by

$$\tilde{Q}(u) = X_{(j)} \ , \ \frac{j-1}{n} < u \leq \frac{j}{n} \ , \ j=1,\ldots,n \ , \tag{1.4}$$

and, for a given set of $k < n$ percentile points $U = \{u_1,\ldots,u_k\}$ satisfying $0 < u_1 <\cdots< u_k < 1$, let $\underline{\tilde{Q}}_U$ and $\underline{Q}_U(\underline{\theta})$ denote the $k \times 1$ vectors

$$\underline{\tilde{Q}}_U = (\tilde{Q}(u_1),\ldots,\tilde{Q}(u_k))'$$

and

$$\underline{Q}_U(\underline{\theta}) = (Q(u_1;\underline{\theta}),\ldots,Q(u_k;\underline{\theta}))'.$$

Then, we propose estimating $\underline{\theta}$ by any vector that minimizes the quadratic form

$$E(\underline{\theta}) = (\underline{\tilde{Q}}_U - \underline{Q}_U(\underline{\theta}))'W(\underline{\theta})(\underline{\tilde{Q}}_U - \underline{Q}_U(\underline{\theta})) \ , \tag{1.5}$$

as a function of $\underline{\theta}$, where $W(\underline{\theta})$ is a user defined $k \times k$ matrix of weights that may be chosen to depend on $\underline{\theta}$. Thus, (1.5) provides an entire family of estimators, $\hat{\underline{\theta}}_U(W)$, indexed by both U and W. The matrix $W(\underline{\theta})$ may be selected to provide specific types of protection or efficiency properties. An optimal choice for W (in a sense to be defined) that is easy to use is

$$W^*(c) = F_U(c) R_U^{-1} F_U(c), \qquad (1.6)$$

where $R_U = \min(u_i, u_j) - u_i u_j$ and $F_U(c)$ is the $k \times k$ diagonal matrix having ith diagonal element $c(1-u_i)Q(u_i)^{(c-1)/c}$. Since R_U is well known to have a tridiagonal inverse, whose typical row has nonzero entries

$$-(u_i - u_{i-1})^{-1}, \; (u_{i+1} - u_{i-1})/\{(u_{i+1} - u_i)(u_1 - u_{i-1})\}$$

$$-(u_{i+1} - u_i)^{-1},$$

the elements of (1.6) can be simply and efficiently evaluated.

Although in general $\hat{\underline{\theta}}_U(W)$ will not have a closed form, the estimator is readily computed using any standard minimization routine such as the IMSL routine ZXMIN or a linearization type procedure (see LaRiccia and Wehrly 1981). It should also be noted that, provided the $\tilde{Q}(u_i)$'s are selected from the uncensored portions of the data, this estimation technique requires no modifications for use with left, right or both left and right censored samples. This is in contrast to most other estimation procedures such as maximum likelihood.

As with all minimum distance estimation procedures, the distance measure provides a natural measure of goodness-of-fit. Under the assumptions of Theorem 1, in the next section, it is easily shown that $nE(\hat{\underline{\theta}}_U(W^*))$ has an asymptotic central chi-squared distribution under the null hypothesis that the data derive from the model (1.1). Thus the goodness-of-fit measure $nE(\hat{\underline{\theta}}_U(W^*))$ could be combined with plots of, for example, $(u, \tilde{Q}(u)-Q(u;\hat{\underline{\theta}}_U(W^*)))$ to ascertain the adequacy of and/or suggest modifications to the Weibull model.

The estimator $\hat{\underline{\theta}}_U(W)$ utilizes only a subset of the entire set of n sample quantiles. Such estimators have received considerable attention in the statistical literature as illustrated by numerous techniques suggested for location and scale parameter estimation (see Harter 1971 and Cheng 1975 for references). This is undoubtedly due, in part, to the fact that they incorporate data compression into the estimation scheme (Eisenberger and Posner 1965) and can provide cost savings when the actual collection of, or reading taken from, an observation is expensive. In addition, using a quantile subset can be advantageous when, for privacy or other reasons, only specific percentiles of the data have been made available for study or, as a further example, in life testing where the savings in time or fraction of items destroyed can be quite large. In some cases, it may be possible to choose the set of percentiles U and, consequently, a technique for the optimal selection of U would be desirable. We consider this problem, and derive an approximate method for its solution, in Section 3.

The next section contains a brief summary of the asymptotic behavior and distribution theory of $\hat{\underline{\theta}}_U(W)$. In Section 3 a technique is presented for selecting U along with consideration of some of the asymptotic efficiency properties of $\hat{\underline{\theta}}_U(W^*)$.

2. Asymptotic Properties of the Estimator

In this section the asymptotic properties of the k-quantile estimator of $\underline{\theta}$ obtained using (1.5) will be presented. Since these results follow from the general theory of k-quantile minimum quantile distance estimators developed by LaRiccia and Wehrly (1982) proofs are omitted. The interested reader is referred to LaRiccia and Wehrly (1982) for the necessary details.

Define the functions

$$f_1(u;c) = c(1-u)Q(u)^{(c-1)/c} \tag{2.1}$$

$$f_2(u;c) = f_1(u;c)Q(u)^{1/c} \tag{2.2}$$

$$f_3(u;c) = -c^{-2}f_1(u;c)Q(u)^{1/c}\ln Q(u) \tag{2.3}$$

and let $B_U(c)$ denote the $k \times 3$ matrix

$$B_U(c) = \{f_j(u_i;c)\}. \tag{2.4}$$

Also let $D(\sigma)$ denote the 3×3 diagonal matrix

$$D(\sigma) = \text{diag }(1,1,\sigma) \tag{2.5}$$

and observe that, using this notation, $F_U(c)$ in (1.6) can be written as

$$F_U(c) = \text{diag}(f_1(u_1;c),\ldots,f_1(u_k;c)). \tag{2.6}$$

We then have the following result concerning the asymptotic behavior and distribution of $\hat{\underline{\theta}}_U(W)$.

Theorem 1. Let $\underline{\theta}_0 = (\mu_0,\sigma_0,c_0)'$ denote the true unknown value of $\underline{\theta}$ and assume that $W(\underline{\theta})$ is positive definite and has elements which possess continuous second partial derivatives, with respect to $\underline{\theta}$, for all $\underline{\theta}$ in some open neighborhood of $\underline{\theta}_0$. Then,

i) as $n\to\infty$ there exists, with probability one, a unique estimator $\hat{\underline{\theta}}_U(W)$ that locally minimizes $E(\underline{\theta})$,

ii) $\hat{\underline{\theta}}_U(W)$ is a consistent estimator of $\underline{\theta}_0$ and $\sqrt{n}(\hat{\underline{\theta}}_U(W)-\underline{\theta}_0)$ converges in distribution to a 3-variate normal distribution with zero mean and variance - covariance matrix

$$V_U(\underline{\theta}_0) = A_U(\underline{\theta}_0)R_U A_U(\underline{\theta}_0)'/\sigma^2 \tag{2.7}$$

where R_U is defined as in (1.6),

$$A_U(\underline{\theta}_0) = [B_U(\underline{\theta}_0)'W(\underline{\theta}_0)B_U(\underline{\theta}_0)]^{-1} \times B_U(\underline{\theta}_0)'W(\underline{\theta}_0)'F_U(c_0)^{-1}$$

and

$$B_U(\underline{\theta}_0) = F_U(c_0)^{-1}B_U(c_0)D(\sigma_0).$$

Regarding the choice (1.6) for $W(\underline{\theta})$ it is possible to show the following optimality property.

Theorem 2. Let $W^*(c)$ be defined as in (1.6) and let $\hat{\underline{\theta}}_U(W^*)$ denote the corresponding estimator of $\underline{\theta}_0$ obtained from (1.5). Then, $\hat{\underline{\theta}}_U(W^*)$ has asymptotic variance - covariance matrix

$$V_U^*(\underline{\theta}_0) = [D(\sigma_0)B_U(c_0)'R_U^{-1}B_U(c_0)D(\sigma_0)]^{-1}/\sigma^2 \tag{2.8}$$

and is optimal in the sense that, for $c > 2$, $1/\det(V_U^*(\underline{\theta}_0))$ is the Fisher information for $\underline{\theta}_0$ in the case that only the order statistics $\tilde{\underline{Q}}_U$ are observed.

3. Selection of U

While the estimator $\hat{\underline{\theta}}_U(W^*)$ is optimal for any given U, its variance - covariance matrix is a function of the specific quantiles selected. In fact, as we shall demonstrate, the placement of the percentile points, u_i, can have a drastic effect on the estimator's efficiency. Consequently, in this section an approximate technique is provided for selecting U. Throughout we take W to be W* defined in (1.6) and, therefore, adopt the notation $\hat{\underline{\theta}}_U(W^*) = \hat{\underline{\theta}}^*_U$. Since the results in this section require the use of the Fisher information matrix for the Weibull we need also assume that c > 2.

Let $I(\underline{\theta})$ denote the Fisher information matrix for the three parameter Weibull distribution which, by reference to Kübler (1979), is seen to be

$$I(\underline{\theta}) = D(\sigma)I(c)D(\sigma)/\sigma^2 \tag{3.1}$$

where

$$I(c) = \begin{bmatrix} (c-1)^2\Gamma(h_2) & c(c-1)\Gamma(h_1) & -(c-1)\Gamma(h_1)H_2/c \\ c(c-1)\Gamma(h_1) & c^2 & -\psi(2) \\ -(c-1)\Gamma(h_1)H_2/c & -\psi(2) & H_1/c^2 \end{bmatrix}, \tag{3.2}$$

Γ and ψ are, respectively, the gamma and digamma functions, $h_i = 1 - i/c$, $i=1,2$,

$$H_1 = \psi'(1) + \psi(2)^2$$

and

$$H_2 = \psi(h_1) + 1.$$

Thus, one method of evaluating the asymptotic relative efficiency of $\hat{\underline{\theta}}^*_U$, for a given U, is to examine the ratio

$$\mathrm{DARE}(\hat{\underline{\theta}}^*_U) \equiv \frac{\det(I(\underline{\theta}_0)^{-1})}{\det(V^*_U(\underline{\theta}_0))} = \frac{\det(B_U(c_0)'R_U^{-1}B_U(c_0))}{\det(I(c_0))}. \tag{3.3}$$

This is seen to be a function of c_0 alone and hence, for a given value of c_0, could be maximized as a function of U to determine optimal percentile points. Of course c_0 is unknown so that, in practice, it is not possible to obtain a U that maximizes (3.3). However, this does suggest an approximate procedure for selecting U, namely, utilize a priori information to provide an initial guess regarding the value of c_0, $\overline{c}$ say, and then maximize

$$\det(B_U(\overline{c})'R_U^{-1}B_U(\overline{c}))/\det(I(\overline{c})) \tag{3.4}$$

with respect to U. It will be seen that, if the U's are chosen as outlined in this section, considerable freedom is allowed in the specification of $\overline{c}$ with little difference in DARE compared to that obtained using c_0. Consequently, this approach will usually work quite well if the experimenter can merely specify a range of possible values for c_0. Alternatively, one could utilize a simple initial estimate for c_0, e.g., the one proposed by Dubey (1967), in situations where such a two-stage procedure is feasible.

Obtaining a value (or values) of U that maximizes (3.4) is a nonlinear optimization problem that is quite difficult in general. However, a value of U that works well in practice can be obtained by first observing that $[B_U(\overline{c})'R_U^{-1}B_U(\overline{c})]^{-1}$ is also the variance - covariance matrix for the best linear unbiased estimator of the parameters of the regression model

$$Y(u_i) = \beta_1 f_1(u_i;\overline{c})+\beta_2 f_2(u_i;\overline{c})+\beta_3 f_3(u_i;\overline{c})+X(u_i),\ i=1,\ldots,k,$$

where the $X(u_i)$'s are zero mean random variables satisfying

$$E[X(u_i)X(u_j)] = \min(u_i,u_j) - u_iu_j.$$

Thus, the theory of regression design in the presence of correlated errors can be applied to the problem of maximizing $\det(B_U(\bar{c})'R_U^{-1}B_U(\bar{c}))^{-1}$ and, hence, (3.4). From results in Sacks and Ylvisaker (1968) (c.f. the Corollary on pg. 62) we obtain an approximate procedure applicable to the DARE criterion wherein the elements of U are chosen as the (k+1)-tiles of the density

$$h(u;\bar{c}) = [\underline{\psi}(u;\bar{c})'I^{-1}(\bar{c})\underline{\psi}(u;\bar{c})]^{1/3} \div \int_0^1 [\underline{\psi}(s;\bar{c})'I^{-1}(\bar{c})\underline{\psi}(s;\bar{c})]^{1/3}ds, \qquad (3.5)$$

and $\underline{\psi}(u;\bar{c})$ denotes the vector with ith component $\frac{d^2}{du^2} f_i(u;\bar{c})$, i=1,2,3 (see also Eubank 1981 for another application of this method of percentile selection). This approach is easily implemented by first using a Gaussian quadrature rule to tabulate the c.d.f. corresponding to $h(u;\bar{c})$ and then obtaining the values of the u_i's by interpolation. Such a program is available upon request.

To ascertain the sensitivity of our procedure for selecting U to the specification of $\bar{c}$ we have computed the ratios $\mathrm{DARE}(\hat{\underline{\theta}}^*_{U(\bar{c})}|c_0) = \det(B'_{U(\bar{c})}(c_0)R_U^{-1}B_{U(\bar{c})}(c_0)) \div \det(I(c_0))$, where $U(\bar{c})$ is the set of percentile points obtained from (3.5), for various values of $\bar{c}$ and c_0. This is tantamount to examining the behavior of the asymptotic relative Fisher efficiency of our estimator based on the percentile points $U(\bar{c})$ when c_0 is the true value for the shape parameter. Table 1 provides a summary of these comparisons for a typical case, k=9,

Table 1. Values of DARE($\hat{\theta}^*_{\underline{U}(\bar{c})} | c_0$) for Various Values of $\bar{c}$ and c_0.

c_0 \ $\bar{c}$	2.5	3.0	3.5	4.0	4.5	5.0	5.5	6.0
2.5	.3136	.3161	.3155	.3144	.3133	.3122	.3114	.3109
3.0	.4518	.4551	.4545	.4532	.4519	.4507	.4498	.4494
3.5	.5238	.5273	.5268	.5256	.5244	.5232	.5224	.5222
4.0	.5662	.5697	.5693	.5682	.5671	.5660	.5653	.5652
4.5	.5933	.5968	.5965	.5955	.5945	.5935	.5929	.5929
5.0	.6119	.6154	.6151	.6142	.6133	.6123	.6118	.6120
5.5	.6252	.6287	.6285	.6277	.6268	.6259	.6254	.6257
6.0	.6352	.6386	.6385	.6377	.6369	.6361	.6356	.6359
6.5	.6429	.6463	.6462	.6455	.6447	.6439	.6435	.6439
7.0	.6490	.6523	.6522	.6516	.6509	.6501	.6498	.6502
7.5	.6539	.6572	.6571	.6566	.6558	.6551	.6548	.6553
8.0	.6579	.6612	.6612	.6606	.6599	.6593	.6590	.6595
8.5	.6613	.6646	.6645	.6640	.6634	.6627	.6625	.6630
9.0	.6641	.6674	.6674	.6669	.6663	.6656	.6654	.6659
9.5	.6666	.6698	.6698	.6693	.6687	.6681	.6679	.6685
10.0	.6687	.6719	.6719	.6715	.6709	.6703	.6701	.6707

Table 1. (continued)

c_0 \ $\bar{c}$	6.5	7.0	7.5	8.0	8.5	9.0	9.5	10.0
2.5	.3104	.3100	.3095	.3092	.3088	.3085	.3082	.3079
3.0	.4490	.4486	.4482	.4479	.4475	.4472	.4469	.4466
3.5	.5219	.5216	.5213	.5210	.5207	.5205	.5202	.5200
4.0	.5651	.5649	.5646	.6544	.5642	.5640	.5638	.5636
4.5	.5929	.5928	.5926	.5925	.5923	.5921	.5920	.5918
5.0	.6120	.6119	.6119	.6118	.6116	.6115	.6114	.6113
5.5	.6258	.6258	.6258	.6257	.6256	.6255	.6254	.6253
6.0	.6361	.6362	.6362	.6362	.6361	.6360	.6360	.6359
6.5	.6441	.6442	.6442	.6442	.6442	.6442	.6441	.6441
7.0	.6504	.6506	.6506	.6507	.6507	.6506	.6506	.6506
7.5	.6555	.6557	.6558	.6559	.6559	.6559	.6558	.6558
8.0	.6598	.6600	.6601	.6601	.6602	.6602	.6602	.6602
8.5	.6633	.6635	.6636	.6637	.6638	.6638	.6638	.6638
9.0	.6663	.6665	.6667	.6668	.6668	.6669	.6669	.6669
9.5	.6668	.6691	.6693	.6694	.6694	.6695	.6695	.6695
10.0	.6711	.6713	.6715	.6716	.6717	.6718	.6718	.6718

and for $c_0, \bar{c}=2.5(.5)10$. As can be seen from examination of the table, the asymptotic efficiencies exhibit no appreciable change for $\bar{c}$ reasonably close to c_0 provided both are sufficiently removed from two, e.g., $c_0, \bar{c} > 3$. Consequently, the precise choice of $\bar{c}$ is not critical provided it is in the vicinity of c_0. One must be cautious in interpreting the results of Table 1, however. They indicate that DARE($\hat{\underline{\theta}}^*_U$) is somewhat insensitive to the choice of U _provided U is chosen as one of the sets obtained using_ (3.5). Such statements may or may not be valid for U's obtained in some other manner and, in fact, it is easy to construct U's for which the converse holds. To illustrate this point we have computed the DARE values for the naive choice of uniformly spaced u_i's when k=9 and c=2.5(.5)10. These DARE's are presented in Table 2 and are seen to be quite low thereby illustrating the improvement to be obtained by using (3.5) in the selection of U.

Table 2. Values of DARE ($\hat{\underline{\theta}}^*_U$) for $U=\{i/10; i=1,\ldots,9\}$ and Various Values of c.

c	DARE($\hat{\underline{\theta}}^*_U$)
2.5	.0409
3.0	.0692
3.5	.0887
4.0	.1028
4.5	.1132
5.0	.1212
5.5	.1274
6.0	.1325
6.5	.1366
7.0	.1401
7.5	.1430
8.0	.1455
8.5	.1476
9.0	.1495
9.5	.1512
10.0	.1526

To conclude it should be noted that, although this paper has concentrated on parameter estimation for the three-parameter Weibull distribution, the techniques developed here can be extended to other distributions and parameter estimation situations as well (provided they satisfy the general conditions presented in LaRiccia and Wehrly (1981) for consistency and asymptotic normality of the minimum quantile distance estimator). For example, it can be shown that analogous statements to those in Sections 1 and 2 hold for the three-parameter lognormal distribution. Thus, one would have little difficulty in obtaining results, such as those in Section 3, for the case of a lognormal model.

REFERENCES

Cheng, S. W. (1975): A unified approach to choosing optimal quantiles for the ABLUE's, Journal of the American Statistical Association 70, 155-159.

Dubey, S. D. (1967): Some percentile estimators for Weibull parameters, Technometrics 9, 119-129.

Eisenberger, I.; Posner, E. C. (1965): Systematic statistics used for data compression in space telemetry, Journal of the American Statistical Association 60, 97-133.

Eubank, R. L. (1981): A density-quantile function approach to optimal spacing selection, Annals of Statistics 9, 494-500.

Harter, H. L. (1971): Some optimizing problems in parameter estimation, In: Optimizing Methods in Statistics (J. S. Rustagi, ed.), Academic Press, New York, 33-62.

Johnson, N. L.; Kotz, S. (1970): Distributions in Statistics: Continuous Univariate Distributions-1, Houghton Mifflin Company, Boston.

Kübler, H. (1979): On the fitting of the three parameter distributions lognormal, gamma and Weibull, Statistiche Hefte 20, 68-125.

LaRiccia, V. N. (1982): Asymptotic properties of weighted L^2 quantile distance estimators, Annals of Statistics 10, 621-624.

LaRiccia, V. N.; Wehrly, T. E. (1982): Asymptotic properties of a family of minimum quantile distance estimators, Technical Report, University of Nebraska-Lincoln.

Mann, N. R.; Schafer, R. E.; Singpurwalla, N. D. (1974): Methods for Statistical Analysis of Reliability and Life Data, New York, Wiley.

Parr, W. C.; Schucany, W. R. (1980). Minimum distance and robust estimation. Journal of the American Statistical Association 75, 616-624.

Sacks, J. and Ylvisaker, D. (1968): Designs for regression problems with correlated errors; many parameters, Annals of Mathematical Statistics 39, 49-69.

T. J. Carmody* and R. L. Eubank*
Department of Statistics
Southern Methodist University
Dallas, Texas 75275
USA

V. N. LaRiccia
Department of Mathematical Sciences
University of Delaware
Newark, Delaware 19711
USA

*Research supported by the Office of Naval Research under contract number N00014-82-K-0207.