Statistical Science
2003, Vol. 18, No. 2, 231–240

Utilizing a Quantile Function Approach to Obtain Exact Bootstrap Solutions

Michael D. Ernst and Alan D. Hutson

Abstract. The popularity of the bootstrap is due in part to its wide applicability and the ease of implementing resampling procedures on modern computers. But careful reading of Efron (1979) will show that at its heart, the bootstrap is a "plug-in" procedure that involves calculating a functional $\theta(\hat{F})$ from an estimate of the c.d.f. F. Resampling becomes invaluable when, as is often the case, $\theta(\hat{F})$ cannot be calculated explicitly. We discuss some situations where working with the sample quantile function, $\hat{Q}$, rather than $\hat{F}$, can lead to explicit (exact) solutions to $\theta(\hat{F})$.

Key words and phrases: Censored data, confidence band, L-estimator, Monte Carlo, order statistics.

1. INTRODUCTION

Nonparametric bootstrap estimation in the i.i.d. continuous univariate setting is traditionally defined as a technique for estimating the quantity $\theta(F)$, where F denotes the distribution function of the random variables $X_1, \dots, X_n$. The estimation consists of "plugging in" the empirical distribution function estimator $\hat{F}(x) = \sum_{i=1}^{n} I_{(X_i \le x)}/n$, where $I_{(.)}$ denotes the indicator function, in place of F in the population quantity $\theta(F)$. For example, the bootstrap estimator of $\theta(F) = E(X_i|F) = \int x\, dF$ is $\theta(\hat{F}) = E(X_i|\hat{F}) = E_{\hat{F}}(X_i) = \int x\, d\hat{F} = \bar{X}$. However, apparent difficulties arise when closed-form expressions based on $\theta(\hat{F})$ are not easily calculated. Approximate bootstrap solutions to complicated problems may be handled using the well-known bootstrap resampling approach described in many bootstrap texts such as those by Efron and Tibshirani (1993), Davison and Hinkley (1997) and Shao and Tu (1995). The appeal of the resampling approach has always been that it provides practitioners with a powerful and easy-to-use tool for tackling difficult, and what might otherwise be intractable analytic problems under more relaxed assumptions than parametric methods. However, resampling techniques introduce additional error through the choice of the Monte Carlo scheme, are sometimes as difficult to implement as the analytical solutions themselves and are vulnerable to manipulation through multiple runs of a computer program.

Other researchers have focused their efforts on using different estimators of F from which to resample in order to compensate for the discreteness of the classic empirical estimator $\hat{F}$ in small samples. For example, Silverman and Young (1987) suggest modifying the resampling procedure by employing a smoothed version of the empirical distribution function. Jiménez Gamero, Muñoz García and Muñoz Reyes (1998) suggest eliminating "outlier" bootstrap samples in the resampling procedure in order to minimize their impact on the estimation procedure. Hutson (1999) employs fractional order statistics to improve the coverage probabilities of analytically calculated bootstrap percentile confidence intervals for quantiles. Lee (1994) has proposed combining parametric and nonparametric bootstrapping procedures by defining an estimator of the form $\varepsilon\theta(\hat{F}) + (1-\varepsilon)\theta(F_{\hat{\phi}})$ that is a convex combination of a nonparametric and para-

Michael D. Ernst is Assistant Professor, Department of Mathematical Sciences, Indiana University–Purdue University Indianapolis, 402 North Blackford Street, LD 270, Indianapolis, Indiana 46202-3216 (e-mail: mernst@math.iupui.edu). Alan D. Hutson is Associate Professor, Department of Biostatistics, University at Buffalo, 249 Farber Hall, 3435 Main Street, Buffalo, New York 14214-3000 (e-mail: ahutson@buffalo.edu).

metric estimator. In practice, Lee suggests that the optimal ε can be estimated from the data. Even though Lee's approach has been shown to be asymptotically optimal based on an MSE criteria, it is oftentimes complicated to carry out in terms of a resampling method. It is, however, an idea worth further exploration.

An alternative approach to examining and modifying the bootstrap resampling procedure is to consider various quantile function estimators. Let $X_{1:n} \le X_{2:n} \le \cdots \le X_{n:n}$ denote the order statistics from an i.i.d. sample of size n from a continuous distribution F. Hutson and Ernst (2000) have illustrated that a nonparametric bootstrap replication can be generated by taking a random sample of size n from a uniform(0, 1) distribution and applying the sample quantile function

$$\hat{Q}(u) = \hat{F}^{-1}(u) = X_{\lfloor nu \rfloor + 1:n} \tag{1}$$

to each uniform random variable, where $\lfloor \cdot \rfloor$ denotes the floor function. By employing $\hat{Q}(u)$, they show that various bootstrap-estimated moments of L-estimators that are not easily calculated by considering $\hat{F}(x)$ may be obtained directly, thus eliminating the resampling step in the estimation altogether. This general quantile function approach has been fruitful for solving a variety of difficult analytical bootstrap problems. One place where this method has paid dividends is for bootstrap problems that deal with censored data. Also, direct calculation can eliminate the need for an outer bootstrap resampling loop in a double bootstrap, making a more efficient algorithm. Section 2 details the exact bootstrap moment calculations for L-estimators, while Section 3 illustrates the calculation of exact bootstrap percentiles. Section 4 outlines the application to censored data, and Section 5 discusses exact bootstrap confidence bands for quantile functions.

2. L-ESTIMATORS

An L-estimator (or L-statistic) is defined as

$$L_n = \sum_{i=1}^{n} c_i X_{i:n},$$

where the choice of constants c_i determines the properties and functionality of L_n. Special cases of note include the mean, the trimmed mean, the median, quick estimators of location and scale and the first and third quartiles. It is possible to give analytic expressions for the exact bootstrap mean and variance estimates for the general L-estimator. The basic approach follows by "plugging in" the quantile function estimator (1) into the well-known expressions for the moments of order statistics (David, 1981). Specifically, the exact nonparametric bootstrap estimators for

$$\begin{aligned} \mu_{r:n} &= E(X_{r:n}), \\ \sigma_{r:n}^2 &= \mathrm{Var}(X_{r:n}), \\ \sigma_{rs:n} &= \mathrm{Cov}(X_{r:n}, X_{s:n}), \end{aligned}$$

are directly estimable and are denoted by $\hat{\mu}_{r:n}$, $\hat{\sigma}_{r:n}^2$ and $\hat{\sigma}_{rs:n}$, respectively. The bootstrap mean and variance of any L_n is simply a linear combination of these estimates. Hutson and Ernst (2000) show that

$$\hat{\mu}_{r:n} = E_{\hat{Q}}(X_{r:n}) = \sum_{j=1}^{n} w_{j(r)} X_{j:n}, \tag{2}$$

$$\hat{\sigma}_{r:n}^2 = \mathrm{Var}_{\hat{Q}}(X_{r:n}) = \sum_{j=1}^{n} w_{j(r)} (X_{j:n} - \hat{\mu}_{r:n})^2, \tag{3}$$

$$\begin{aligned} \hat{\sigma}_{rs:n} &= \mathrm{Cov}_{\hat{Q}}(X_{r:n}, X_{s:n}) \\ &= \sum_{j=2}^{n} \sum_{i=1}^{j-1} w_{ij(rs)} (X_{i:n} - \hat{\mu}_{r:n})(X_{j:n} - \hat{\mu}_{s:n}) \\ &\quad + \sum_{j=1}^{n} v_{j(rs)} (X_{j:n} - \hat{\mu}_{r:n})(X_{j:n} - \hat{\mu}_{s:n}), \end{aligned} \tag{4}$$

where the weights $w_{j(r)}$, $w_{ij(rs)}$ and $v_{j(rs)}$ are given in Appendix A.

Using the formulation presented above, we can illustrate the relative error of the resampling method for estimating the bootstrap standard error of an L-estimator based on a small to medium number of bootstrap replications. This is illustrated through the use of a classical dataset described by Stigler (1977) in his study of the performance of robust estimators applied to "real" data. It is interesting to note that Stigler (1977) concluded that a slightly trimmed mean is the best estimator of location and that overall the 10% trimmed mean emerged as the "recommended estimator."

In the absence of the exact bootstrap variance estimator given above, the practitioner has been left with various rules of thumb in choosing the number of resamples needed for an accurate estimate of the bootstrap standard error. Efron and Tibshirani (1993, page 52) state that "Very seldom are more than $B = 200$ replications needed for estimating a standard error" and that "$B = 50$ is often enough to give a good estimate." Booth and Sarkar (1998) examine this question and provide general recommendations of 800 resamples in order to obtain accurate standard error estimates. The exact bootstrap variance estimator eliminates the resampling error as well as the need for arbitrarily choosing a value for B.

Table 1 contains a small dataset from Stigler (1977). These measurements are James Short's 1763 determinations of the parallax of the sun (in seconds of a degree) based on the 1761 transit of Venus (see Stigler's appendix for details). For these data, we consider three common location L-estimators: the 10% trimmed mean the 25% trimmed mean, and the median. Their exact bootstrap standard error estimates are 0.167, 0.165 and 0.165, respectively.

Figure 1 shows the ratio of 200 estimates of the bootstrap standard error to the exact bootstrap standard error for each of the three L-estimators. Each of the resampling estimates is based on $B = 100$ resamples. The resampling error is as much as 20% for the trimmed means and sometimes exceeds that for the median. When the number of resamples is increased to $B = 500$ (Figure 2) and $B = 1000$ (Figure 3), the error is reduced, but can still be as high as 10% for the trimmed means and 15% for the median.

The absolute error from resampling becomes magnified when the standard error estimate is used in interval calculations. For example, in a 95% confidence interval, the absolute error is increased by nearly a factor of 4. This illustrates the point made by Efron and Tibshirani (1993, page 52) that "Much bigger values of B are required for bootstrap confidence intervals." This resampling error is eliminated entirely by the use of the exact expressions for the bootstrap mean and variance of an L-estimator given in this section.

TABLE 1
Short's 1763 determinations of the parallax of the sun (in seconds of a degree) based on the 1761 transit of Venus (Stigler, 1977) ($n = 18$)

8.50	8.50	7.33	8.64	9.27	9.06
9.25	9.09	8.50	8.06	8.43	8.44
8.14	7.68	10.34	8.07	8.36	9.71

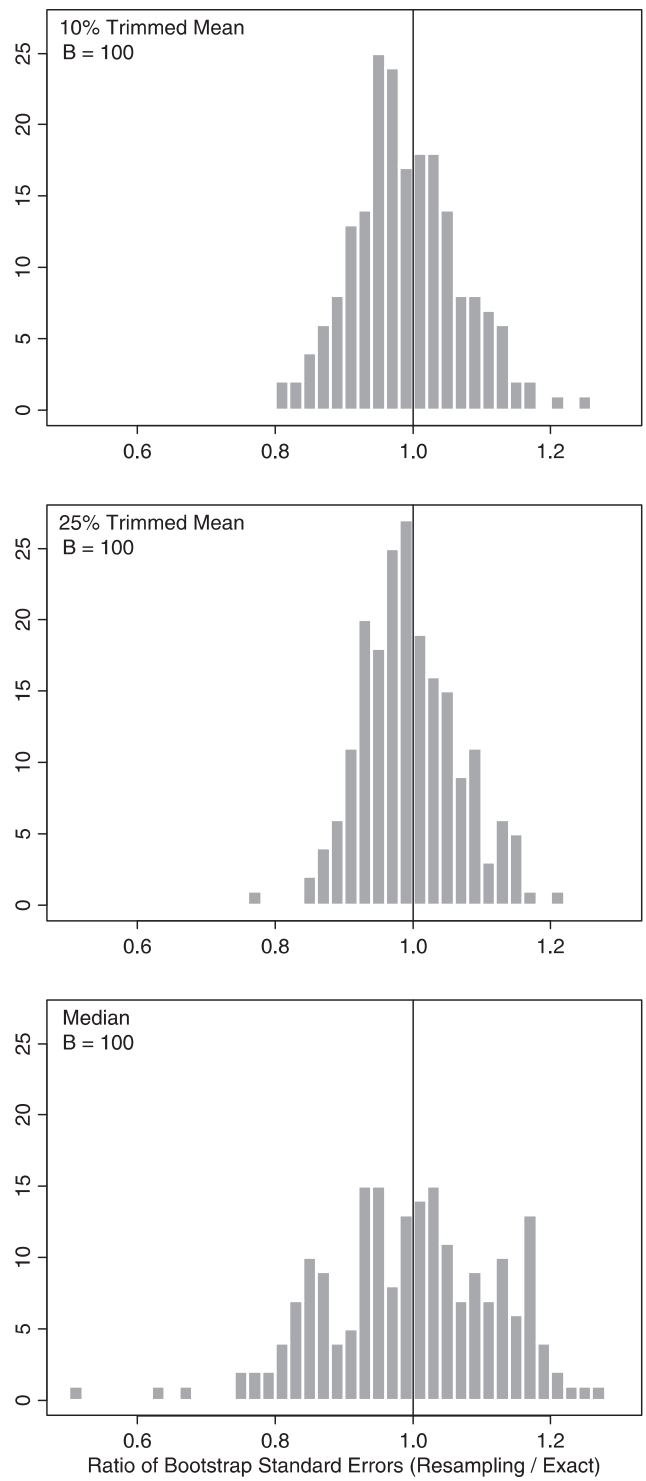

FIG. 1. *Ratio of 200 bootstrap standard errors, based on $B = 100$ resamples each, to the exact bootstrap standard error of the 10% trimmed mean, 25% trimmed mean and median of the data in Table 1.*

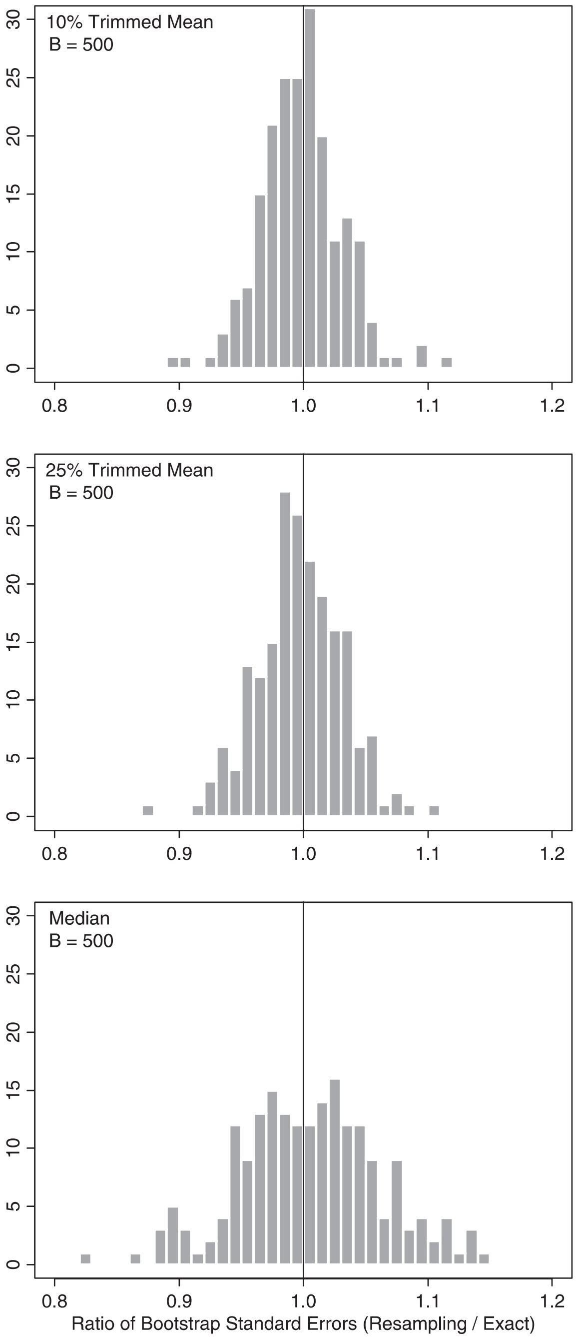

FIG. 2. *Ratio of* 200 *bootstrap standard errors, based on* $B = 500$ *resamples each, to the exact bootstrap standard error of the* 10% *trimmed mean,* 25% *trimmed mean and median of the data in Table* 1.

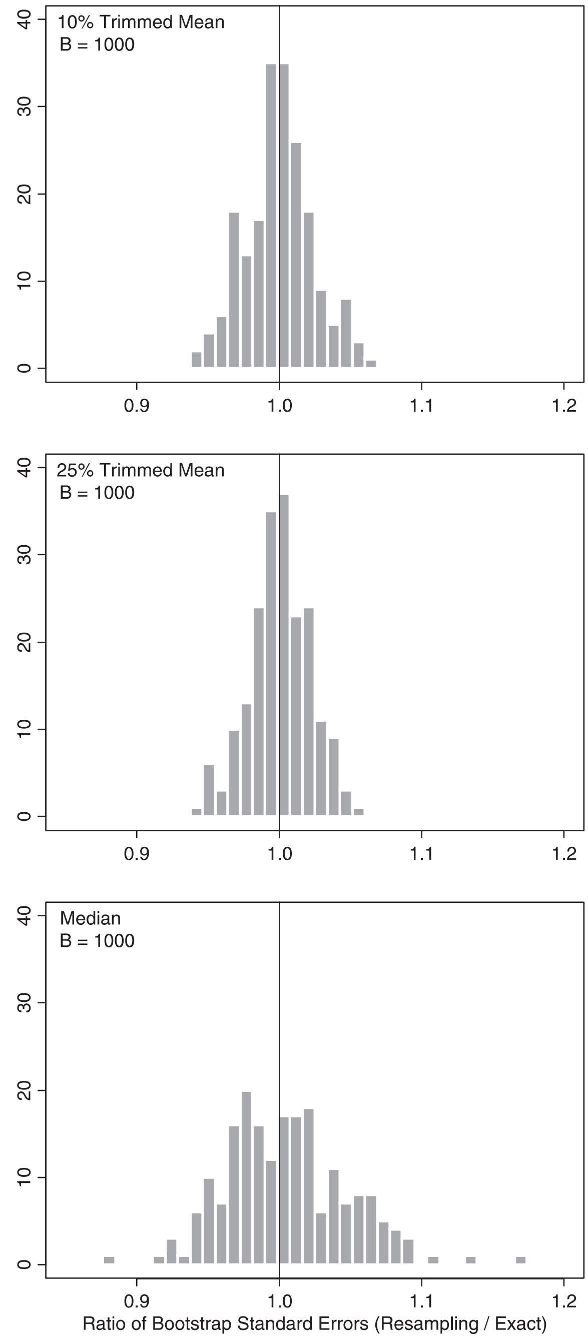

FIG. 3. *Ratio of* 200 *bootstrap standard errors, based on* $B = 1000$ *resamples each, to the exact bootstrap standard error of the* 10% *trimmed mean,* 25% *trimmed mean and median of the data in Table* 1.

3. EXACT PERCENTILE CONFIDENCE INTERVALS

Using a straightforward minimization algorithm, we can obtain the exact nonparametric bootstrap percentiles for any statistic

$$T_n = T(X_{j_1:n}, X_{j_2:n}, \ldots, X_{j_k:n}) \tag{5}$$

that is a function of k order statistics, where $1 \le j_1 < j_2 < \cdots < j_k \le n$ and $1 \le k \le n$. The only case where the exact percentiles have been demonstrated to have analytical solutions is the case where $T_n = X_{i:n}$ is exactly equal to a specific order statistic (David, 1981).

The method for calculating the exact bootstrap percentiles for T_n follows from a version of the

TABLE 2
Patients' baseline urinary apABG (nmol/d) ($n = 24$)

67.9	7.1	14.0	10.9	3.1	8.5	646.3	0.5	6.2	9.4	10.3	4.9
136.0	138.5	297.7	184.3	10.6	433.5	275.7	3.3	230.8	12.0	7.8	21.4

definition of the α-quantile of the random variable Y given by the value of θ satisfying

$$\inf_{\theta\in\mathbb{R}} E\left\{\frac{|Y-\theta|+(2\alpha-1)(Y-\theta)}{2} - \frac{|Y|+(2\alpha-1)Y}{2}\right\}. \tag{6}$$

The properties of (6) are outlined in Abdous and Theodorescu (1992) who generalize the definition of the α-quantile to $\mathbb{R}^k$. Most notably, the assumption that $E(Y) < \infty$ is not a necessary condition for the existence of the α-quantile.

The 100α bootstrap percentile of T_n, denoted by the quantile function estimator $\hat{Q}_{T_n}(\alpha)$, is obtained through the exact bootstrap estimate of the linear component of (6) given by

$$E\{|T_n-\theta|+(2\alpha-1)(T_n-\theta)\}, \tag{7}$$

and then minimizing the estimate of (7) with respect to θ. The other linear components of (6) do not factor into the minimization procedure and can be ignored. The general process is outlined in Appendix B. For the specific case of L-estimators, the process is greatly simplified.

EXAMPLE. We use the data in Table 2 from Caudill, Gregory, Hutson and Bailey (1998) to illustrate the behavior of the bootstrap percentiles based on resampling compared to the exact bootstrap percentiles. These data are measurements of the urinary folate catabolite, acetamidobenzolyglutamate (apABG), used to assess folate requirements in both pregnant and non-pregnant women. Table 3 shows the exact 2.5% and 97.5% bootstrap percentiles for the median, trimean and interquartile range (IQR) for the data in Table 2. The resample approximations of these percentiles are also given based on $B = 100$, 500 and 1000 resamples. We can see from Table 3 that the lower and upper limits of the simulated percentile intervals vary quite a bit over the replication sizes (B) and that it is not quite clear when convergence is obtained.

4. EXTENDING EXACT METHODS TO CENSORED DATA

The previous methods can be extended to provide some exact bootstrap quantities useful in survival analysis by making use of the well-known product-limit estimator introduced by Kaplan and Meier (1958), with the empirical survival function estimator defined as

$$\hat{S}(t) = \begin{cases} \prod_{T_{(j)}\le t}\left(\frac{n-j}{n-j+1}\right)^{\delta_{(j)}}, & t < T_{(n)}, \\ 0, & t \ge T_{(n)}, \end{cases} \tag{8}$$

where $T_{(1)} \le T_{(2)} \le \cdots \le T_{(n)}$ are the order statistics corresponding to the i.i.d. sample of n failure or censoring times $T_1, T_2, \dots, T_n$, and $\delta_{(1)}, \delta_{(2)}, \dots, \delta_{(n)}$ are censoring indicators corresponding to the ordered T_i's, respectively. A value of $\delta_{(i)} = 1$ indicates that $T_{(i)}$ is uncensored, while a value of $\delta_{(i)} = 0$ indicates that $T_{(i)}$ is censored. No analytical nonparametric bootstrap estimators for $\theta(F)$ have been developed based on the Kaplan–Meier estimator or any other nonparametric estimator, accounting for censored data.

The typical Monte Carlo approach for censored data is similar to the classical approach described above

TABLE 3
Exact and approximate 95% *bootstrap percentile intervals for the data in Table* 2

Statistic	T_n	**Exact**	$B = 100$	$B = 500$	$B = 1000$
Median	11.45	(8.50,136.00)	(9.40,136.00)	(8.15,136.00)	(8.50,136.00)
Trimean*	54.03	(10.60,144.38)	(9.75,139.58)	(11.33,140.75)	(10.43,144.20)
IQR**	176.50	(9.10,289.91)	(6.20,289.90)	(9.10,289.90)	(7.80,290.60)

*Tukey's trimean defined as $\hat{Q}(1/4)/4 + \hat{Q}(1/2)/2 + \hat{Q}(3/4)/4$.
**Interquartile range defined as $\hat{Q}(3/4) - \hat{Q}(1/4)$.

where the pairs $(T_{(i)}, \delta_{(i)})$, $i = 1, 2, \dots, n$, are sampled with replacement with probability $1/n$ assigned to each pair. The bootstrap quantities of interest are then obtained via information summarized over recalculations of the Kaplan–Meier estimator (or other similar estimators) given by each bootstrap replication (see Akritas, 1986, or Barber and Jennison, 1999). This resampling approach has some undesirable properties, which can be avoided through the exact approach. It should be clear that even though $\hat{S}(t)$ is defined to be a proper estimator of the survivor function, estimation problems can arise when a given bootstrap resample has a large proportion of censored observations. Other approaches are outlined in Doss and Gill (1992).

4.1 Moment Estimators of Survival Quantities

The general approach for estimating the functional $\theta(F)$ with censored data is straightforward. We may directly calculate the expressions for important survival quantities such as $E(T^k)$, $E(\hat{S}(t)^k)$, $E([\hat{S}^{-1}(1-u)]^k)$ and $\mathrm{Cov}(\hat{S}^{-1}(1-u), \hat{S}^{-1}(1-v))$, corresponding to the lifetime central moments, the survival fraction moments, the survival quantile moments and the covariance between survival quantiles, respectively.

For an i.i.d. continuous sample of size n from a distribution having positive support, let $T_i = \min(X_i, C_i)$, where X_i denotes a failure time and C_i denotes a right censoring time. The value of X_i is known only if $X_i \le C_i$. In addition, let the indicator variable $\delta_{(i)} = 1$ if $X_i \le C_i$, and 0 otherwise, and let $X_{(1)} \le X_{(2)} \le \cdots \le X_{(m)}$ denote the ordered observed failure times, $m \le n$. Then the exact kth bootstrap moment about the origin for the random variable T is given by

$$E_{\hat{S}}(T^k) = \begin{cases} \displaystyle\sum_{i=1}^{m} X_{(i)}^k\big(\hat{S}(X_{(i-1)}) - \hat{S}(X_{(i)})\big), & \delta_{(n)} = 1, \\ \displaystyle\sum_{i=1}^{m} X_{(i)}^k\big(\hat{S}(X_{(i-1)}) - \hat{S}(X_{(i)})\big) \\ \qquad + T_{(n)}^k \hat{S}(X_{(m)}), & \delta_{(n)} = 0, \end{cases}$$

where $\hat{S}$ is the Kaplan–Meier estimator defined in (8) and by definition $X_{(0)} = 0$ implies $\hat{S}(X_{(0)}) = 1$. It follows that the bootstrap mean and variance of T are given by $\hat{\mu}_T = E_{\hat{S}}(T)$ and $\hat{\sigma}_T^2 = E_{\hat{S}}(T^2) - [E_{\hat{S}}(T)]^2$, respectively.

Define the quantile function estimator $\hat{F}^{-1}(u) = X_{(i)}$ whenever $\hat{S}(X_{(i)}) \le 1-u < \hat{S}(X_{(i-1)})$, $i = 1, 2, \dots, m$. Then it follows that the exact kth bootstrap moment estimator of $\hat{F}^{-1}(u)$ for the case $\delta_{(n)} = 1$ is given by

$$\begin{aligned} &E_{\hat{S}}[(\hat{F}^{-1}(u))^k] \\ &\quad = \sum_{j=1}^{m} X_{(j)}^k[\beta_{p,q}(1 - \hat{S}(X_{(j)})) \\ &\qquad\qquad - \beta_{p,q}(1 - \hat{S}(X_{(j-1)}))], \end{aligned} \tag{9}$$

where $\beta_{p,q}(\cdot)$ denotes the c.d.f. of a beta distribution with parameters $p = \lfloor nu \rfloor + 1$ and $q = n - \lfloor nu \rfloor + 2$. For the case $\delta_{(n)} = 0$, add the term $T_{(n)}^k \beta_{p,q}(1 - \hat{S}(X_{(m)}))$ to (9).

As a by-product, (9) can be modified to provide a kernel quantile function estimator, which reduces to the kernel quantile function estimator of Harrell and Davis (1982) when all observations are uncensored. Unlike the more traditional kernel quantile function estimators with symmetric kernel functions and having the form $h_n^{-1} \int_0^1 \hat{F}^{-1}(t) K((t-u)/h_n)\,dt$, the extension of the Harrell–Davis estimator has an asymmetric kernel. See Sheather and Marron (1990) for a general study of kernel quantile function estimators based on uncensored observations and Padgett (1986) for the development of kernel quantile function estimators for censored data. The extension of the Harrell–Davis kernel quantile function estimator for censored observations is defined as

$$\begin{aligned} \hat{F}_K^{-1}(u) = \sum_{j=1}^{m} X_{(j)}&[\beta_{p,q}(1 - \hat{S}(X_{(j)})) \\ &- \beta_{p,q}(1 - \hat{S}(X_{(j-1)}))], \end{aligned}$$

where $\beta_{p,q}(\cdot)$ denotes the c.d.f. of a beta distribution with fractional parameters $p = nu$ and $q = n(1-u)$. Xiang (1995) provides a theoretical examination of the strengths and weaknesses of these types of estimators.

The exact bootstrap estimator of $\mathrm{Cov}(\hat{F}^{-1}(u), \hat{F}^{-1}(v))$ for the case $\delta_{(n)} = 1$ is calculated using a similar approach to that found in Hutson and Ernst (2000). Other quantities such as the first two exact bootstrap moment estimators of the random variable $\hat{S}(t)$ also follow; see Hutson (2002a).

5. EXACT CONFIDENCE BANDS FOR QUANTILE FUNCTIONS

For calculating confidence bands for quantile functions, Doss and Gill (1992) suggest a bootstrapping method for censored or uncensored data, which is carried out under less strict regularity conditions than those of Csörgő (1983). Their approach consists of determining an estimated critical value $\hat{d}_\alpha$ by bootstrapping the quantity

$$\sup_{0<u<1}\left|\frac{\hat{Q}(u)-Q(u)}{\hat{v}^{(i)}(u)}\right|, \tag{10}$$

such that $\hat{d}_\alpha$ is the 100α percentile of (10) obtained by replacing $\hat{Q}(u)-Q(u)$ with $\hat{Q}^*(u)-\hat{Q}(u)$ in the numerator, where $\hat{Q}^*(u)$ is the quantile function estimated from one bootstrap replication. The confidence band is then defined as

$$\big(\hat{Q}(u)-\hat{d}_\alpha\hat{v}^{(i)}(u),\ \hat{Q}(u)+\hat{d}_\alpha\hat{v}^{(i)}(u)\big), \tag{11}$$

where $\hat{v}^{(i)}(u)$ is an arbitrarily chosen scale estimator of $\hat{Q}(u)$. The suggested choice for $\hat{v}^{(i)}(u)$ is the interquantile range. This confidence band is basically the bootstrap version of the approach described in Csörgő (1983). Doss and Gill (1992) suggest two layers of bootstrapping: first, the bootstrap distribution of $\hat{v}^{(i)}(u)$ needs to be determined, followed by the bootstrap distribution of (10). Note, however, as described in Section 2, exact estimates for $\hat{v}^{(i)}(u)$ are available, eliminating the outer bootstrap resampling layer. A negative feature of the Doss–Gill method is the fact that neither the lower nor the upper bands are constrained to be monotone even though the underlying function is known to be monotone. Other alternative bootstrapping methods are given by Breth (1980), Li, Hollander, McKeague, and Yang (1996) and Yeh (1996).

Alternatively, Hutson (2002b) shows that one can obtain exact bootstrap confidence bands based on Steck's determinant. Let $U_{i:n}$ denote the corresponding uniform order statistic such that $X_{i:n}=Q(U_{i:n})$. Steck (1971) has proven that

$$\begin{aligned}&P\big(Q(l_i)\le X_{i:n}\le Q(u_i), i=1,2,\dots,n\big)\\&\quad=\det(t_{ij}),\end{aligned} \tag{12}$$

where $t_{ij}=\binom{j}{j-i+1}(u_i-l_j)_+^{j-i+1}$ or 0 as $j-i+1$ is nonnegative or negative across $i=1,2,\dots,n$ and $j=1,2,\dots,n$, and $(x)_+=\max(0,x)$. An exact bootstrap confidence interval can be obtained via (12) by first solving for a single constant (denoted c_α) in the following expression based on the joint c.d.f. of the order statistics,

$$\begin{aligned}&F_{X_{1:n},X_{2:n},\dots,X_{n:n}}\bigg(Q_{X_{1:n}}(0),\ Q_{X_{2:n}}\Big(1-\frac{\alpha}{2c_\alpha}\Big),\\&\qquad\qquad\dots,\ Q_{X_{n:n}}\Big(1-\frac{\alpha}{2c_\alpha}\Big)\bigg)\\&\quad-F_{X_{1:n},X_{2:n},\dots,X_{n:n}}\bigg(Q_{X_{1:n}}\Big(\frac{\alpha}{2c_\alpha}\Big),\ Q_{X_{2:n}}\Big(\frac{\alpha}{2c_\alpha}\Big),\\&\qquad\qquad\dots,\ Q_{X_{n:n}}(1)\bigg)\\&\quad=1-\alpha,\end{aligned} \tag{13}$$

where $1-\alpha$ is the desired confidence level and $Q_{X_{i:n}}(\cdot)$ denotes the quantile function for the ith order statistic. The exact bootstrapping component of the estimation follows once c_α has been determined simply by "plugging in" the corresponding estimates of Q and F. Hutson (2002b) utilizes the linear interpolation estimators, $\tilde{Q}$ and $\tilde{F}$, for Q and F in place of the step function estimators in order to improve the accuracy in small samples.

The $(1-\alpha)100\%$ confidence band for $Q(u)$, $1/(n+1)<u<n/(n+1)$, consists of the set of points {lower, upper} given by

$$\begin{aligned}&\bigg\{\bigg(-\infty,\ \tilde{Q}_{X_{2:n}}\Big(1-\frac{\alpha}{2c_\alpha}\Big),\ \tilde{Q}_{X_{3:n}}\Big(1-\frac{\alpha}{2c_\alpha}\Big),\\&\qquad\qquad\dots,\ \tilde{Q}_{X_{n:n}}\Big(1-\frac{\alpha}{2c_\alpha}\Big)\bigg),\\&\quad\bigg(\tilde{Q}_{X_{1:n}}\Big(\frac{\alpha}{2c_\alpha}\Big),\ \tilde{Q}_{X_{2:n}}\Big(\frac{\alpha}{2c_\alpha}\Big),\\&\qquad\qquad\dots,\ \tilde{Q}_{X_{n-1:n-1}}\Big(\frac{\alpha}{2c_\alpha}\Big),+\infty\bigg)\bigg\},\end{aligned}$$

corresponding to the knots $i/(n+1)$ for large samples, $i=1,2,\dots,n$, where $\tilde{Q}_{X_{i:n}}(\cdot)=\tilde{Q}(Q_{U_{i:n}}(\cdot))$. The points corresponding to the upper band are joined by the ceiling function, while the points on the lower band are joined by a floor function. This choice of the step functions helps maintain the overall confidence level of $1-\alpha$.

This confidence band is essentially an estimated transformation of a confidence band for a uniform quantile function. The transformation guarantees that the upper and lower bands will be monotone; this is consistent with the true (and estimated) quantile function. In fact, if more information is known about the

TABLE 4
Empirical coverage probabilities for $n = 10$, 20 *and* 30

	Coverage probabilities ($1-\alpha = 0.95$)								
	$n = 10$			$n = 20$			$n = 30$		
Distribution	**EB**	**CR**	**DG***	**EB**	**CR**	**DG***	**EB**	**CR**	**DG***
Normal	0.95	0.92	0.95	0.95	0.93	0.94	0.95	0.94	0.95
Exponential	0.95	0.93	0.98	0.94	0.93	0.97	0.94	0.95	0.97
Uniform	0.94	0.92	0.95	0.94	0.94	0.95	0.94	0.95	0.96
Logistic	0.95	0.92	0.95	0.95	0.93	0.94	0.95	0.95	0.94
Extreme value	0.95	0.92	0.96	0.94	0.93	0.96	0.95	0.95	0.95
Laplace	0.95	0.93	0.95	0.96	0.94	0.95	0.95	0.95	0.93

true quantile function, this can be easily incorporated into the confidence band. For example, the lower band will be nonnegative for a quantile function of a nonnegative random variable.

Hutson (2002b) compared the exact bootstrap (EB) method with the method of Csörgő and Révész (CR) and the resampling bootstrap method of Doss and Gill (DG). They were compared with respect to their 95% coverage probabilities when sampling is from a standard normal distribution, exponential distribution, uniform distribution, logistic distribution, extreme value distribution and Laplace distribution, for samples of size $n = 10$, 20 and 30. The results, contained in Table 4, are based on 10,000 simulations for the EB and CR methods. Only 1000 simulations were used for the DG method since the simulation study for this method is very computationally intensive. In addition, for the DG method 1000 resamples within each simulation were needed in order to estimate d_α. Since the exact bootstrap interquartile range of $\tilde{Q}(u)$ is given by

$$\widetilde{IQR}_u = \tilde{Q}(Q_\beta(3/4)) - \tilde{Q}(Q_\beta(1/4)),$$

where $Q_\beta(\cdot)$ denotes the $\beta(\lfloor nu \rfloor, \lfloor n(1-u) \rfloor)$ quantile function, it was used in order to estimate the scale, $\hat{v}^{(i)}(u)$, in (11). Using the exact estimate for $\hat{v}^{(i)}(u)$ eliminates the need for double bootstrapping suggested by Doss and Gill (1992).

In Table 4, we see that the CR and EB methods have very good coverage rates for small samples with the EB method having slightly better and precise overall coverage rates for samples of size $n = 10$. In our original simulation study, the DG method had a very poor coverage rate. Even though Doss and Gill (1992) illustrated the asymptotic correctness of their approach, care needs to be taken with small samples. Specific problems occur in the tail regions due to the discreteness of the quantile function estimator. Furthermore, the choice of step function used in the estimation does make a difference in the coverage rates for small samples. We repeated the simulation making two minor heuristic adjustments: (1) we limited the range of u to $1/n \le u \le (n-2)/n$, and (2) we defined the upper bound to be a ceiling function and the lower bound to be a floor function. We see from Table 4 that this modified method, denoted DG*, performs fairly well.

Besides having good coverage rates, as seen in Table 4, the EB method also tends to cover a larger range of the quantile function. We illustrate this by comparing the EB and CR methods on a dataset described by Stigler (1977) in his study of the performance of robust estimators. The data contained in Table 5 are the results of a 1798 experiment by Cavendish who set about trying to determine the mean density of the earth. The "true" value is 5.517 g/ccm. The EB (solid line) and CR (dashed line) 95% confidence bands are given in Figure 4 overlayed with the estimated parametric quantile function $\bar{x} + s\Phi^{-1}(u)$ applied under the normality assumption. We see that the EB and CR bands are similar in the middle of the distribution, while the EB method provides a greater range of coverage on both ends of the lower and the upper bands.

TABLE 5
Cavendish's 1798 *determinations of the density of the earth* (*g/ccm*) *relative to water* (*Stigler*, 1977) ($n = 29$)

5.50	5.61	4.88	5.07	5.26	5.55	5.36	5.29	5.58	5.65
5.57	5.53	5.62	5.29	5.44	5.34	5.79	5.10	5.27	5.39
5.42	5.47	5.63	5.34	5.46	5.30	5.75	5.68	5.85	

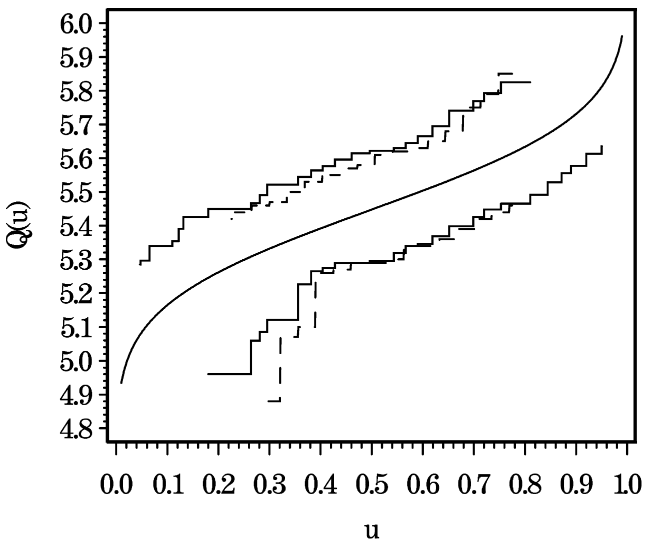

FIG. 4. *95% confidence bands for $Q(u)$ using the EB (solid line) and CR (dashed line) methods from the data in Table* 5.

APPENDIX A

The weights used in (2), (3) and (4) for the exact bootstrap moments of an L-estimator are given by

$$w_{j(r)} = r\binom{n}{r}\left[B\left(\frac{j}{n}; r, n-r+1\right) - B\left(\frac{j-1}{n}; r, n-r+1\right)\right],$$

$$\begin{aligned} w_{ij(rs)} = {}_nC_{rs} \sum_{k=0}^{s-r-1} \binom{s-r-1}{k} \frac{(-1)^{s-r-1-k}}{s-k-1} &\cdot \left[\left(\frac{i}{n}\right)^{s-k-1} - \left(\frac{i-1}{n}\right)^{s-k-1}\right] \\ &\cdot \left[B\left(\frac{j}{n}; k+1, n-s+1\right) - B\left(\frac{j-1}{n}; k+1, n-s+1\right)\right] \end{aligned}$$

and

$$\begin{aligned} v_{j(rs)} = {}_nC_{rs} \sum_{k=0}^{s-r-1} \binom{s-r-1}{k} \frac{(-1)^{s-r-1-k}}{s-k-1} &\cdot \left\{B\left(\frac{j}{n}; s, n-s+1\right) - B\left(\frac{j-1}{n}; s, n-s+1\right) \right. \\ &- \left(\frac{j-1}{n}\right)^{s-k-1} \cdot \left[B\left(\frac{j}{n}; k+1, n-s+1\right) \right. \\ &\left.\left. - B\left(\frac{j-1}{n}; k+1, n-s+1\right)\right]\right\}, \end{aligned}$$

where ${}_nC_{rs} = n!/[(r-1)!(s-r-1)!(n-s)!]$ and $B(x; a, b) = \int_0^x t^{a-1}(1-t)^{b-1}\,dt$ is the incomplete beta function.

APPENDIX B

To derive the exact bootstrap percentiles of (5), first let

$$g(u_1, u_2, \ldots, u_k) = n! \prod_{i=0}^{k} \frac{(u_{i+1} - u_i)^{j_{i+1}-j_i-1}}{(j_{i+1} - j_i - 1)!}$$

be the joint density of the k uniform order statistics from a sample of size n, $U_{j_1:n}, U_{j_2:n}, \ldots, U_{j_k:n}$, where $u_{k+1} = 1$, $u_0 = 0$, $j_{k+1} = n+1$ and $j_0 = 0$. The exact bootstrap estimate of (7) is given by

$$\begin{aligned} &\sum_{i_k=1}^{n} \sum_{i_{k-1}=1}^{i_k} \cdots \sum_{i_1=1}^{i_2} w_{i_1 \cdots i_{k-1} i_k (j_1 \cdots j_{k-1} j_k)} \\ &\quad \cdot \big[|T(X_{i_1:n}, \ldots, X_{i_k:n}) - \theta| + (2\alpha - 1) \cdot \big(T(X_{i_1:n}, \ldots, X_{i_k:n}) - \theta\big)\big], \end{aligned} \tag{14}$$

where the weights are given by

$$w_{i_1 \cdots i_{k-1} i_k (j_1 \cdots j_{k-1} j_k)} = \int_{(i_k-1)/n}^{i_k/n} \int_{(i_{k-1}-1)/n}^{I_{i_{k-1}}} \cdots \int_{(i_1-1)/n}^{I_{i_1}} g(u_1, \ldots, u_k)\,du_1 \cdots du_k$$

and

$$I_{i_a} = \begin{cases} u_{a+1}, & \text{if } i_a = i_{a+1}, \\ i_a/n, & \text{otherwise.} \end{cases}$$

The exact bootstrap percentiles of (5) are given by the value of θ that minimizes (14).

Expression (14) follows from the re-expression of (7) using the quantile function $Q(u)$. Specifically, if h is the joint p.d.f. of $X_{j_1:n}, X_{j_2:n}, \dots, X_{j_k:n}$, then

$$
\begin{aligned}
&E\{|T_n-\theta|+(2\alpha-1)(T_n-\theta)\}\\
&\quad=\int_{-\infty}^{\infty}\int_{-\infty}^{x_k}\cdots\int_{-\infty}^{x_2}[|T_n-\theta|\\
&\qquad\qquad+(2\alpha-1)(T_n-\theta)]\\
&\qquad\qquad\times h(x_1,\dots,x_k)\,dx_1\cdots dx_k\\
&\quad=\int_0^1\int_0^{u_k}\cdots\int_0^{u_2}\{|T(Q(u_1),\dots,Q(u_k))-\theta|\\
&\qquad\qquad+(2\alpha-1)[T(Q(u_1),\dots,\\
&\qquad\qquad Q(u_k))-\theta]\}\\
&\qquad\qquad\times g(u_1,\dots,u_k)\,du_1\cdots du_k.
\end{aligned}
\tag{15}
$$

The result follows by substituting $\hat{Q}(u)$ given by (1) into (15) and noting that $\hat{Q}(u)$ is constant in the region given by $(i-1)/n \le u < i/n$, $i = 1, 2, \dots, n$.

ACKNOWLEDGMENT

The authors thank Professor Govind S. Mudholkar for bringing this topic to their attention.

REFERENCES

ABDOUS, B. and THEODORESCU, R. (1992). Note on the spatial quantile of a random vector. *Statist. Probab. Lett.* **13** 333–336.

AKRITAS, M. (1986). Bootstrapping the Kaplan–Meier estimator. *J. Amer. Statist. Assoc.* **81** 1032–1038.

BARBER, S. and JENNISON, C. (1999). Symmetric tests and confidence intervals for survival probabilities and quantiles of censored survival data. *Biometrics* **55** 430–436.

BOOTH, J. G. and SARKAR, S. (1998). Monte Carlo approximation of bootstrap variances. *Amer. Statist.* **52** 354–357.

BRETH, M. (1980). Quantile estimation and probability coverages. *Austral. J. Statist.* **22** 207–211.

CAUDILL, M. A., GREGORY, J. F., HUTSON, A. D. and BAILEY, L. B. (1998). Folate catabolism in pregnant and nonpregnant women with controlled folate intakes. *J. Nutrition* **128** 204–208.

CSÖRGŐ, M. (1983). *Quantile Processes with Statistical Applications*. SIAM, Philadelphia.

DAVID, H. A. (1981). *Order Statistics*, 2nd ed. Wiley, New York.

DAVISON, A. C. and HINKLEY, D. V. (1997). *Bootstrap Methods and Their Application*. Cambridge Univ. Press.

DOSS, H. and GILL, R. D. (1992). An elementary approach to weak convergence for quantile processes with applications to censored survival data. *J. Amer. Statist. Assoc.* **87** 869–877.

EFRON, B. (1979). Bootstrap methods: Another look at the jackknife. *Ann. Statist.* **7** 1–26.

EFRON, B. AND TIBSHIRANI, R. (1993). *An Introduction to the Bootstrap*. Chapman and Hall, New York.

HARRELL, F. E. and DAVIS, C. E. (1982). A new distribution-free quantile estimator. *Biometrika* **69** 635–640.

HUTSON, A. D. (1999). Calculating nonparametric confidence intervals for quantiles using fractional order statistics. *J. Appl. Statist.* **26** 343–353.

HUTSON, A. D. (2002a). Analytical bootstrap methods for censored data. *J. Appl. Math. Decis. Sci.* **6** 129–141.

HUTSON, A. D. (2002b). "Exact" bootstrap confidence bands for the quantile function via Steck's determinant. *J. Comput. Graph. Statist.* **11** 471–482.

HUTSON, A. D. and ERNST, M. D. (2000). The exact bootstrap mean and variance of an L-estimator. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **62** 89–94.

JIMÉNEZ GAMERO, M. D., MUÑOZ GARCÍA, J. and MUÑOZ REYES, A. (1998). Bootstrapping statistical functionals. *Statist. Probab. Lett.* **39** 229–236.

KAPLAN, E. L. and MEIER, P. (1958). Nonparametric estimation from incomplete observations. *J. Amer. Statist. Assoc.* **52** 457–481.

LEE, S. M. S. (1994). Optimal choice between parametric and non-parametric bootstrap estimates. *Math. Proc. Cambridge Philos. Soc.* **115** 335–363.

LI, G., HOLLANDER, M., MCKEAGUE, I. W. and YANG, J. (1996). Nonparametric likelihood ratio confidence bands for quantile functions from incomplete survival data. *Ann. Statist.* **24** 628–640.

PADGETT, W. J. (1986). A kernel-type estimator of a quantile function from right-censored data. *J. Amer. Statist. Assoc.* **81** 215–222.

SHAO, J. and TU, D. (1995). *The Jackknife and Bootstrap*. Springer, New York.

SHEATHER, S. J. and MARRON, J. S. (1990). Kernel quantile estimators. *J. Amer. Statist. Assoc.* **85** 410–416.

SILVERMAN, B. W. and YOUNG, G. A. (1987). The bootstrap: To smooth or not to smooth? *Biometrika* **75** 469–479.

STECK, G. P. (1971). Rectangle probabilities for uniform order statistics and the probability that the empirical distribution function lies between two distribution functions. *Ann. Math. Statist.* **42** 1–11.

STIGLER, S. M. (1977). Do robust estimators work with *real* data? (with discussion). *Ann. Statist.* **5** 1055–1098.

XIANG, X. (1995). Deficiency of the sample quantile estimator with respect to kernel quantile estimators for censored data. *Ann. Statist.* **23** 836–854.

YEH, A. B. (1996). Bootstrap percentile confidence bands based on the concept of curve depth. *Comm. Statist. Simulation Comput.* **25** 905–922.