Department of Statistics

University of Wisconsin
Madison, WI 53706

Technical Report No. 728

November 1983

ESTIMATING AN ENDPOINT OF A DISTRIBUTION WITH RESAMPLING METHODS

(ABBREVIATED TITLE: ESTIMATING AN ENDPOINT)

by

Wei-Yin Loh

Department of Statistics

University of Wisconsin, Madison

Estimating an endpoint of a distribution with resampling methods

(Abbreviated title: Estimating an endpoint)

by Wei-Yin Loh

Department of Statistics

University of Wisconsin, Madison

Summary

The problem of estimating an endpoint of a distribution is revisited, using the bootstrap and random subsample methods. Contrary to an example in Bickel and Freedman (1981) suggesting that these methods do not work here, it is shown that one can in fact construct asymptotically valid confidence intervals. However, the results also indicate that resampling methods may be more model-dependent than originally thought.

AMS 1980 subject classifications. 62G05, 62G15.

Key words and phrases. Bootstrap, random subsampling, confidence intervals.

Introduction. Bickel and Freedman (1981) give the following as a counter-example to Efron's (1979) bootstrap method. Let $X_1, X_2, \ldots, X_n$ be independent, identically distributed random variables from the uniform distribution F on the interval $(0,\theta)$. Using the natural pivot $n(\theta - X_{(n)})/\theta$, where $X_{(i)}$ denotes the ith order statistic, they observe that (i) $n(\theta - X_{(n)})/\theta$ tends to a limiting exponential distribution, and (ii) with probability one, the conditional distribution of the bootstrap quantity $n(X_{(n)} - X^*_{(n)})/X_{(n)}$ does not have a weak limit. Here $(X^*_1, \ldots, X^*_n)$ denotes a bootstrap sample. Since the bootstrap distribution does not approximate the true distribution of the pivot well even in the limit, these authors conclude that the bootstrap method does not work for this situation.

In this paper we re-examine the problem more generally for any F with a right endpoint θ and belonging to the domain of attraction of the type II extreme value law, i.e. we only assume that there is $\delta \geq 0$ such that (cf. Gnedenko, 1943)

$$(1.1) \qquad \lim_{x \to 0-} \{1-F(cx+\theta)\}/\{1-F(x+\theta)\} = c^{\delta} \qquad \text{for all } c > 0.$$

The uniform, as well as any distribution with a finite, non-zero density at θ, corresponds to $\delta = 1$. It is easy to verify that under (1.1), the above observations generally hold true, namely, (i)' $n^{1/\delta}(\theta - X_{(n)})$ tends to a limiting distribution, and, (ii)' with probability one, the conditional distribution of $n^{1/\delta}(X_{(n)} - X^*_{(n)})$ does not have any weak limit. However, using a method of constructing bootstrap intervals first considered but apparently abandoned by Efron (1979), we show that these intervals are asymptotically valid for precisely one value of $\delta \neq 1$. This result has two implications. The first is that the bootstrap method is more "robust" than first thought, since it can provide valid inferences even without the bootstrap distribution being close to the true distribution of the pivot. On the other hand, since the method works for only one value of δ, the result suggests that it is highly model-dependent.

Even more surprisingly, it will be shown that if we repeat the whole argument with Hartigan's (1969) random subsampling method instead of the bootstrap, then (i)' and (ii)' again hold with $X^*_{(n)}$ replaced with the largest value in each random subsample. But now if F is the uniform distribution, the random subsample intervals have exact coverage probabilities for all sample sizes! In fact these intervals are asymptotically valid for all F with $\delta = 1$.

The non-uniform performance of the bootstrap method can be corrected if we knew δ. This is accomplished by using a "generalized" bootstrap, which resamples the observations with unequal probabilities depending on δ. The validity of the resulting intervals is proved in Section 4.

The twin problem of finding point estimates of θ is also considered. Essentially our procedure is to derive median- and mean-bias corrected estimates based on $X_{(n)}$ using the bootstrap and random subsample distributions. Two of the estimates so obtained turn out to have been proposed earlier in Robson and Whitlock (1964) and Cooke (1979). A third is new. This application of the bootstrap method does not appear to have attracted much attention in the bootstrap literature.

2. Survey of known results.

One of the first to consider the problem of estimating an endpoint of a distribution from a nonparametric viewpoint is Miller (1964), whose main purpose was to show that Tukey's (1958) original suggestion for constructing jackknife t-intervals can give incorrect answers. This and subsequent papers assume the following framework. Let $X_1,\ldots,X_n$ be independent, identically distributed random variables from a distribution $F(x-\theta)$ such that $F(x) < 1$ for $x < 0$ and $F(0) = 1$, for some finite θ. Further, $F(x)$ is assumed to belong to the domain of attraction of a type II extreme value law, i.e. it satisfies (1.1) for some $\delta \geq 0$.

Miller (1964) showed that Quenouille's (1949) jackknife estimate based on the naive estimate $X_{(n)}$ is

$$\hat{\theta}_J = X_{(n)} + n^{-1}(n-1)(X_{(n)}-X_{(n-1)}). \tag{2.1}$$

When θ is a truncation parameter, i.e. $\delta = 1$, the particular bias expansion of $X_{(n)}$ led Robson and Whitlock (1964) to propose the modified jackknife estimate

$$\hat{\theta}_{RW} = 2X_{(n)} - X_{(n-1)}, \tag{2.2}$$

which is asymptotically equivalent to (2.1).

Noting that neither of these estimators has smaller asymptotic mean squared error than $X_{(n)}$ when $\delta = 1$, Cooke (1979) obtained the estimator

$$(2.3) \qquad 2X_{(n)} - \sum_{i=0}^{n-1} [(1-i/n)^n - \{1-(i+1)/n\}^n]X_{(n-i)},$$

which is asymptotically equivalent to

$$(2.4) \qquad \hat{\theta}_c = 2X_{(n)} - (1-e^{-1}) \sum_{i=0}^{n-1} e^{-i} X_{(n-i)}.$$

He showed that this has smaller asymptotic mean squared error than (2.2) for $\delta = 1$, but not for $\delta > 1$.

Better estimates are available if δ is assumed known. Cooke (1979) gave formulas for the best constants c_1 and c_2 minimizing the mean squared errors within the respective classes

$$X_{(n)} + c_1(X_{(n)}-X_{(n-1)})$$

and

$$X_{(n)} + c_2\{X_{(n)}-(1-e^{-1}) \sum_{i=0}^{n-1} e^{-i} X_{(n-i)}\}.$$

Note that $c_1 = c_2 = 1$ yields (2.2) and (2.4). In another paper, Cooke (1980) also derived the best linear estimator based on a fixed number of the largest order statistics. For solutions assuming conditions stronger than (1.1), see Hall (1982) and the references therein.

When δ is unknown, the choice between (2.1) - (2.4) is in some sense not absolutely critical, since the mean squared error of each is $O(u_n^2)$, where $u_n = F^{-1}(1-n^{-1})$. Thus every estimator is at least consistent. In contrast, the situation with interval estimation is very different. Miller (1964) showed that the jackknife t-intervals give completely wrong coverage probabilities. Robson and Whitlock (1964) obtained the interval

$$(X_{(n)}, X_{(n)} + \alpha^{-1}(1-\alpha)(X_{(n)} - X_{(n-1)}))$$

which has asymptotic coverage probability $1 - \alpha$ only for $\delta = 1$. Cooke (1979) generalized this to

$$(X_{(n)}, X_{(n)} + \{(1-\alpha)^{-\nu}-1\}^{-1}(X_{(n)} - X_{(n-1)})). \tag{2.5}$$

This has asymptotic coverage probability $1 - \alpha$ if and only if (1.1) holds with $\delta = 1/\nu$. Weissman (1981) further generalized (2.5) to the "two-sided" interval involving lower order statistics

$$I_W^i(\nu;p_1,p_2) = (X_{(n)}+r_i(p_1)(X_{(n)}-X_{(n-i)}),\quad X_{(n)} + r_i(p_2)(X_{(n)}-X_{(n-i)})) \tag{2.6}$$

where $r_i(p) = [\{1-(1-p)^{1/i}\}^{-\nu}-1]^{-1}$. He showed that this interval has asymptotic coverage probability $p_2 - p_1$ if (1.1) holds with $\delta = 1/\nu$. Clearly (2.5) is a special case of (2.6).

3. Bootstrap and random subsample procedures.

Let $\hat{\theta}^*$ denote the largest value in a bootstrap or random subsample, and P_* the associated resampling probabilities. It is easy to verify that for $i = 1,2,\ldots,n-1$, the bootstrap distribution is

$$(3.1) \qquad P_*(\hat{\theta}^* \le X_{(n-i)}) = \{(n-i)/n\}^n \to e^{-i},$$

and the random subsample distribution is

$$(3.2) \qquad P_*(\hat{\theta}^* \le X_{(n-i)}) = (2^{n-i}-1)/(2^n-1) \to 2^{-i}.$$

We first prove statements (i)' and (ii)' mentioned in the introduction.

Theorem 3.1. Under (1.1) and as $n \to \infty$,

(i) $n^{1/\delta}(\theta - X_{(n)})$ converges in distribution to that of $Z^{1/\delta}$, where Z is the standard exponential random variable, and

(ii) w.p.1, the conditional distribution of $n^{1/\delta}(X_{(n)} - \hat{\theta}^*)$, under either (3.1) or (3.2), does not have a weak limit.

Proof. From standard results concerning extremal processes (cf. Weissman, 1981), we know that for fixed k, the joint distribution of $(n^{1/\delta}(\theta - X_{(n)}), n^{1/\delta}(\theta - X_{(n-1)}),\ldots,n^{1/\delta}(\theta - X_{(n-k)}))$ converges to that of the random variables $(Z_1^{1/\delta},(Z_1+Z_2)^{1/\delta},\ldots,(Z_1+\ldots Z_{k+1})^{1/\delta})$, where the Z_i's are independent standard exponential random variables. Hence (i) follows. It further follows that for each fixed k, $n^{1/\delta}(X_{(n)} - X_{(n-k)})$ converges in distribution to that of $(Z_1+\ldots+Z_{k+1})^{1/\delta} - Z_1^{1/\delta}$. The Hewitt-Savage zero-one law now implies that $\limsup n^{1/\delta}(X_{(n)}-X_{(n-k)}) = \infty$ and $\liminf n^{1/\delta}(X_{(n)}-X_{(n-k)})=0$ a.s. This together with (3.1) and (3.2) yield part (ii) of the theorem.

We proceed to show that despite this fact, the bootstrap and random subsample methods can still give useful results. First we consider interval estimation of θ. Efron (1981, 1982) has given two methods, called the "percentile" and "bias-corrected percentile" methods, but because both yield intervals contained within the support of the bootstrap distribution, they clearly do not work here. Instead we resurrect another method originally criticised in Efron (1979, Remark D). Let t_α be the 100α percentile of the bootstrap distribution of $\hat{\theta}^*$, i.e.

$$P_*(t_\alpha < \hat{\theta}^* \leq X_{(n)}) = 1-\alpha,$$

or equivalently

$$P_*(t_\alpha - \hat{\theta} < \hat{\theta}^* - \hat{\theta}) = 1-\alpha, \tag{3.3}$$

since $\hat{\theta} = X_{(n)}$. If we believe, as the bootstrap method would have us believe, that the Monte Carlo distribution of $\hat{\theta}^* - \hat{\theta}$ is close to the true distribution of $\hat{\theta} - \theta$, (3.3) suggests the approximation $P(t_\alpha - \hat{\theta} < \hat{\theta} - \theta) \simeq 1-\alpha$. This gives $(X_{(n)}, 2X_{(n)} - t_\alpha)$ as an approximate $1-\alpha$ one-sided confidence interval for θ. We will show that this approximation is asymptotically valid under (1.1) for some value of δ.

Although Efron (1979) has advocated splitting the bootstrap probabilities at the endpoints of the intervals in other situations, it turns out that because of the asymmetric nature of the present problem, this should not be done here. Thus if $\alpha = e^{-i}$ for some integer i, we deduce from (3.1) that an approximate $1 - \alpha$ bootstrap interval for θ is $(X_{(n)}, 2X_{(n)} - X_{(n-i)})$. This interval however has associated $\alpha = 2^{-i}$ if we use random subsampling (3.2) instead of (3.1). Clearly the two methods cannot both be correct at any one instance. For values of α not of the form e^{-i} or 2^{-i}, but less than e^{-1} or 1/2 respectively, we obtain $1 - \alpha$ intervals by performing the usual trick of randomization.

<u>Theorem 3.2</u>. The bootstrap interval $(X_{(n)}, 2X_{(n)} - X_{(n-i)})$ has asymptotically exact confidence coefficient $1 - \alpha = 1 - e^{-i}$ if and only if (1.1) holds with

$$\delta = \log(1-e^{-1})/\log(.5). \tag{3.4}$$

Similarly, the random subsample interval $(X_{(n)}, 2X_{(n)} - X_{(n-i)})$ has asymptotically exact confidence coefficient $1 - \alpha = 1 - 2^{-i}$ if and only if (1.1) holds with $\delta = 1$.

<u>Proof</u>. Recall from (2.6) that Weissman's (1981) interval $I_W^i(\nu;0,1-\alpha) = (X_{(n)}, X_{(n)} + r_i(1-\alpha)(X_{(n)} - X_{(n-i)}))$ has asymptotic coverage probability $1 - \alpha$ if (1.1) holds with $\delta = 1/\nu$. The theorem follows by equating $r_i(1-\alpha)$ to 1 and solving for δ.

A stronger result obtains if we specialize F to be the uniform distribution.

Theorem 3.3. Let F be the uniform distribution. Then for all $(1/2)^{n-1} \leq \alpha \leq 1/2$, the intervals obtained by random subsampling have exact coverage probabilities for all n.

Proof. Direct calculation shows that

$$P(\theta < 2X_{(n)} - X_{(n-i)}) = 1 - 2^{-i} \quad \text{for } i = 1,2,\ldots,n-1.$$

It is interesting to obtain improved estimates from $X_{(n)}$ by subtracting from it the bootstrap and random subsample estimates of bias. From (3.2) we see that $(X_{(n)} - X_{(n-1)})$ estimates the median-bias of $X_{(n)}$. Therefore a median-bias corrected estimator of θ is $2X_{(n)} - X_{(n-1)}$ which coincides with Robson and Whitlock's estimator given in (2.2). No corresponding estimator is available from (3.1) since the bootstrap distribution puts approximately $1 - e^{-1}$ of its mass on $X_{(n)}$. However, we can use both (3.1) and (3.2) to obtain estimates of the mean-bias in $X_{(n)}$. Subtracting these estimates from $X_{(n)}$ turns out to produce respectively $\hat{\theta}_c$ in (2.4) and a new estimator, $2X_{(n)} - 2^n(2^n-1)^{-1} \sum_{i=0}^{n-1} 2^{-i-1} X_{(n-i)}$, which we can write approximately as

$$\hat{\theta}_{RS} = 2X_{(n)} - (1/2) \sum_{i=0}^{n-1} 2^{-i} X_{(n-i)}. \tag{3.5}$$

Theorem 3.4. Let $u_n = F^{-1}(1-n^{-1})$ and $\nu = 1/\delta$. Then under (1.1) and as $n \to \infty$,

$$(\theta-u_n)^{-1}\ \text{Bias}\ (X_{(n)}) \to \Gamma(\nu+1),$$

$$(\theta-u_n)^{-1}\ \text{Bias}\ (\hat{\theta}_{RW}) \to (\nu-1)\Gamma(\nu+1),$$

$$(\theta-u_n)^{-1}\ \text{Bias}\ (\hat{\theta}_{C}) \to \{(1-e^{-1})^{-\nu}-2\}\Gamma(\nu+1),$$

$$(\theta-u_n)^{-1}\ \text{Bias}\ (\hat{\theta}_{RS}) \to (2^{\nu}-2)\Gamma(\nu+1).$$

Proof. Follows by direct calculation using the formulas in Cooke (1979, section 3).

This theorem shows that both $\hat{\theta}_{RW}$ and $\hat{\theta}_{RS}$ remove the first order bias when $\delta = 1$, and $\hat{\theta}_C$ does the same when δ is given by (3.4). It may be verified that when n is large and $\delta \neq 1$, $\hat{\theta}_{RW}$ and $\hat{\theta}_{RS}$ have biases in the same direction and $\lim_{n\to\infty} |\text{Bias}(\hat{\theta}_{RS})/\text{Bias}(\hat{\theta}_{RW})| \geq 1$.

Theorem 3.5. Let ν and u_n be as in Theorem 3.4 and

$$H(p) = 4\Gamma(2\nu+1) + \Gamma(2\nu+1)(1-p)^2(1-p^2)^{-2\nu-1}$$
$$- 4\Gamma(\nu+1)(1-p) \sum_{i=0}^{\infty} p^i\Gamma(2\nu+i+1)/\Gamma(\nu+i+1)$$
$$+ 2(1-p)^2 \sum_{i=1}^{\infty} p^i\{\Gamma(2\nu+i+1)/\Gamma(\nu+i+1)\} \sum_{j=0}^{i-1} p^j\Gamma(\nu+j+1)/\Gamma(j+1).$$

Then under (1.1) and as $n \to \infty$,

$$(\theta-u_n)^{-2}\text{MSE}(X_{(n)}) \to \Gamma(2\nu+1)$$

$$(\theta-u_n)^{-2}\text{MSE}(\hat{\theta}_{RW}) \to \Gamma(2\nu+1)\{(2\nu^2-\nu+1)/(\nu+1)\}$$

$$(\theta-u_n)^{-2}\text{MSE}(\hat{\theta}_c) \to H(e^{-1})$$

$$(\theta-u_n)^{-2}\text{MSE}(\hat{\theta}_{RS}) \to H(.5).$$

Proof. Again use the formulas in Cooke (1979). Note however that his formula (12) is incorrect.

The following table gives some numerical values for the RHS of the above quantities.

Table 3.1. Some values of $\lim\ (\theta-u_n)^{-2}$ MSE $(\hat{\theta})$ for $\hat{\theta} = X_{(n)}$, $\hat{\theta}_{RW}$, $\hat{\theta}_c$ and $\hat{\theta}_{RS}$

δ	$X_{(n)}$	$\hat{\theta}_{RW}$	$\hat{\theta}_c$	$\hat{\theta}_{RS}$
$\log(1-e^{-1})/\log(1/2)$	6.172	9.969	4.699	7.367
1	2.000	2.000	1.331	1.333
2	1.000	0.667	0.719	0.599
3	0.903	0.602	0.704	0.609
4	0.886	0.620	0.728	0.650
5	0.887	0.651	0.754	0.687

4. A generalized bootstrap.

A heuristic explanation can be given for the peculiar values of δ in Theorem 3.2. From standard theory (cf. Weissman, 1981) we know that for each fixed k,

$$(4.1) \qquad P\{(\theta - X_{(n)})/(X_{(n)} - X_{(n-k)}) < 1\} \to 1 - (1 - .5^{\delta})^k$$

as $n \to \infty$ under (1.1). The bootstrap method approximates the LHS with

$$(4.2) \qquad P_*\{(X_{(n)} - X^*_{(n)})/(X_{(n)} - X_{(n-k)}) < 1\}$$

$$= P_*(X^*_{(n)} > X_{(n-k)}) \to 1 - e^{-k}$$

Equating the RHS of (4.1) and (4.2) yields the value of δ in (3.4). A similar heuristic explanation works for random subsampling.

This suggests that if δ is known, and we draw bootstrap samples in such a way that

$$(4.3) \qquad P_*(X^*_{(n)} > X_{(n-k)}) = 1 - (1 - .5^{\delta})^k,$$

then the resulting intervals will have asymptotically correct coverage probabilities. Given $(X_1, \ldots, X_n)$, let $p_k^{(n)}$ be the probability that $X_{(k)}$ is sampled at each draw of this "generalized" bootstrap.

Theorem 4.1. Suppose that (1.1) holds with δ known. Then the generalized bootstrap yields asymptotically valid intervals for θ if

$$p_n^{(n)} = 1 - (1-.5^\delta)^{1/n},$$

$$\sum_{k=n-j}^{n} p_k^{(n)} = 1 - (1-.5^\delta)^{(j+1)/n}, \quad j = 1,2,\ldots,n-2,$$

and

$$p_1^{(n)} = (1-.5^\delta)^{(n-1)/n}.$$

Proof. The values of $p_k^{(n)}$ clearly satisfy (4.3), and the theorem follows from (4.1) and (4.2).

Corollary 4.1. If (1.1) holds with known δ and

(4.4) $$\alpha = (1-.5^\delta)^k$$

for some k, then the interval $(X_{(n)}, 2X_{(n)}-X_{(n-k)})$ for θ produced by the generalized bootstrap has asymptotic confidence coefficient $1 - \alpha$.

Proof. As for Theorem 3.2.

We now consider estimates derived from the generalized bootstrap. Here the bootstrap estimate of mean bias is

$$.5^\delta \sum_{k=0}^{n-2} (1-.5^\delta)^k X_{(n-k)} + (1-.5^\delta)^{n-1} X_{(1)} - X_{(n)}.$$

This gives as a bias-corrected estimator of θ,

$$\theta_{\text{Boot}}^{(\delta)} = 2X_{(n)} - .5^\delta \sum_{k=0}^{n-2} (1-.5^\delta)^k X_{(n-k)} - (1-.5^\delta)^{n-1} X_{(1)}.$$

Theorem 4.2. Let u_n and $H(\cdot)$ be as defined in Theorems 3.4 and 3.5. Under (1.1),

$$(\theta-u_n)^{-1} \text{ Bias } (\hat{\theta}^{(\delta)}_{\text{Boot}}) \to 0$$

and

$$(\theta-u_n)^{-2} \text{ MSE } (\hat{\theta}^{(\delta)}_{\text{Boot}}) \to H(1-.5^{\delta})$$

as $n \to \infty$.

Proof. Same as for Theorems 3.4 and 3.5.

We note that $\hat{\theta}^{(\delta)}_{\text{Boot}}$ is essentially $\hat{\theta}_{RS}$ with $1 - .5^{\delta}$ substituted for 1/2. The following table gives some values of the asymptotic MSE.

Table 4.1. Values of $H(1-.5^{\delta})$

δ	$\log(1-e^{-1})/\log(.5)$	1	2	3	4	5
$H(1-.5^{\delta})$	4.70	1.33	0.46	0.29	0.21	0.16

A better idea of the efficiency of $\hat{\theta}^{(\delta)}_{\text{Boot}}$ may be had by comparing with Cooke's (1980) results for the best linear estimator $\hat{\theta}^{(\delta)}_{\text{Lin}}(r)$ based on the r largest order statistics. For example $\hat{\theta}^{(1)}_{\text{Boot}}$ has approximately the same MSE as $\hat{\theta}^{(1)}_{\text{Lin}}(3)$, and MSE $(\hat{\theta}^{(3)}_{\text{Boot}}) \simeq$ MSE $(\hat{\theta}^{(3)}_{\text{Lin}}(6))$.

5. Remarks.

We have assumed in the last section that δ is known. If it is unknown, the generalized bootstrap may be made adaptive by replacing δ with a consistent estimate. De Haan (1981) showed that one such estimate is

$$\hat{\delta} = \log m/\log \{(X_{(n-2)}-X_{(m)})/(X_{(n-1)}-X_{(n-2)})\}$$

where $m \to \infty$ and $m/n \to 0$, as $n \to \infty$. Clearly the adaptive version of Theorem 4.1 holds. On the other hand, whether δ is known or estimated, the conclusions in Theorem 3.1 remain true with the generalized bootstrap.

It may be argued that, if δ is unknown, $\theta - X_{(n)}$ is not the right quantity to bootstrap, since its limiting distribution, after standardization, is not independent of δ. This criticism is not entirely valid, because there is nothing in the original formulation of the bootstrap method which requires that only pivotal quantities be bootstrapped. (Recall that if $\bar{X}$ and μ denote the sample and population means respectively, $\bar{X}-\mu$ can be usefully bootstrapped even though it is not an asymptotically pivotal quantity when the population variance is unknown.) Asymptotically distribution-free quantities for the present problem do in fact exist. One such, reported in Weissman (1982), is $\log\{(\theta-X_{(n-m)})/(\theta-X_{(n)})\}/\log\{(\theta-X_{(n-k)})/(\theta-X_{(n-m)})\}$, where $1 \le m < k < n$; see also de Haan (1981) for another. For m and k fixed and $n \to \infty$, this has a limiting distribution, under (1.1), which is independent of θ and δ. Knowledge of this distribution will of course lead to approximate confidence intervals for θ. This is done in

Weissman (1982). Unfortunately, any attempt at bootstrapping the obvious quantity

$$\log\{(X_{(n)}-X^*_{(n-m)})/(X_{(n)}-X^*_{(n)})\}/\log\{(X_{(n)}-X^*_{(n-k)})/(X_{(n)}-X^*_{(n-m)})\}$$

immediately runs into difficulties. This is because the latter is undefined whenever two or more of $X^*_{(n)}$, $X^*_{(n-m)}$ and $X^*_{(n-k)}$ are equal to $X_{(n)}$, an event which occurs with substantial probability for almost all samples $(X_1,\ldots,X_n)$.

REFERENCES

1. Bickel, P. J. and Freedman, D. A. (1981). Some asymptotic theory for the bootstrap. Ann. Statist. 9, 1196-1217.

2. Cooke, P. (1979). Statistical inference for bounds of random variables. Biometrika 66, 367-374.

3. Cooke, P. (1980). Optimal linear estimation of bounds of random variables. Biometrika 67, 257-258.

4. de Haan, L. (1981). Estimation of the minimum of a function using order statistics. J. Amer. Statist. Assoc. 76, 467-469.

5. Efron, B. (1979). Bootstrap methods: another look at the jackknife. Ann. Statist. 7, 1-26.

6. Efron, B. (1981). Nonparametric standard errors and confidence intervals. Canad. J. Statist. 9, 139-172.

7. Efron, B. (1982). The jackknife, the bootstrap and other resampling plans. SIAM, Philadelphia.

8. Gnedenko, B. (1943). Sur la distribution limite du terme maximum d'une série aléatoire. Ann. Math. 44, 423-454.

9. Hartigan, J. A. (1969). Using subsample values as typical values. J. Amer. Statist. Assoc. 69, 383-393.

10. Hall, P. (1982). On estimating the endpoint of a distribution. Ann. Statist. 10, 556-568.

11. Miller, R. G. (1964). A trustworthy jackknife. Ann. Math. Statist. 35, 1594-1605.

12. Quenouille, M. H. (1949). Approximate tests of correlation in time-series. J. R. Statist. Soc. B 11, 68-84.

13. Robson, D. S. and Whitlock, J. H. (1964). Estimation of a truncation point. Biometrika 51, 33-39.

14. Tukey, J. (1958). Bias and confidence in not quite large samples. Abstract, Ann. Math. Statist. 29, 614.

15. Weissman, I. (1981). Confidence intervals for a threshold parameter. Comm. Statist. A10, 549-557.

16. Weissman, I. (1982). Confidence intervals for the threshold parameter II: unknown shape parameter. Comm. Statist. A11, 2451-2474.