COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
AT YALE UNIVERSITY

Box 2125, Yale Station
New Haven, Connecticut 06520

COWLES FOUNDATION DISCUSSION PAPER NO. 1157

A SIMPLE COUNTEREXAMPLE TO THE BOOTSTRAP

Donald W. K. Andrews

August 1997

A Simple Counterexample to the Bootstrap

Donald W. K. Andrews*
Cowles Foundation for Research in Economics
Yale University

June 1997
Revised: July 1997

Abstract

The bootstrap of the maximum likelihood estimator of the mean of a sample of iid normal random variables with mean μ and variance one is not asymptotically correct to first order when the mean is restricted to be nonnegative. The problem occurs when the true value of the mean μ equals zero. This counterexample to the bootstrap generalizes to a wide variety of estimation problems in which the true parameter may be on the boundary of the parameter space. We provide some alternatives to the bootstrap that are asymptotically correct to first order.

We consider two types of bootstrap percentile confidence intervals in the above example. We find that they both have asymptotic coverage probability that exceeds the nominal asymptotic level when the true value of the mean μ equals zero.

1 The Counterexample

The literature contains a number of examples in which the bootstrap of Efron (1979) does not consistently estimate the true distribution of a statistic correctly to first order. The examples in the literature are all nonstandard in some way or other. Here we provide an example that is very simple and quite close to being standard. Furthermore, straightforward generalizations of this example are of importance in many applications.

We consider the maximum likelihood estimator of the mean of a sample of iid normal random variables with mean μ and variance one (denoted $N(\mu, 1)$) when the mean is restricted to be nonnegative. The maximum likelihood estimator in this case is just the maximum of the sample mean and zero. When the true mean is zero, the bootstrap is not asymptotically correct to first order.

Let $\{X_i : i \geq 1\}$ be a sequence of independent identically distributed (iid) $N(\mu, 1)$ random variables. Suppose the parameter space for μ is $R^+ := \{y : y \geq 0\}$. The

*The author gratefully acknowledges the research support of the National Science Foundation via grant number SBR-9410675. The author thanks Peter Bickel and Joe Romano for references.

maximum likelihood estimator of μ in this case is $\widehat{\mu}_n := \max(\overline{X}_n, 0)$, where $\overline{X}_n := \frac{1}{n}\sum_{i=1,\ldots,n} X_i$. It is easy to see that

$$n^{1/2}(\widehat{\mu}_n - \mu) \xrightarrow{d} \begin{cases} Z & \text{if } \mu > 0 \\ \max(Z,0) & \text{if } \mu = 0 \end{cases} \text{ as } n \to \infty, \text{where } Z \sim N(0,1). \quad (1)$$

Let $\{X_i^* : i \leq n\}$ be iid with $X_i^* \sim \widehat{F}_n$, where $\widehat{F}_n(x) := \frac{1}{n}\sum_{i=1,\ldots,n} 1(X_i \leq x)$. The bootstrap maximum likelihood estimator $\widehat{\mu}_n^*$ is defined by $\widehat{\mu}_n^* := \max(\overline{X}_n^*, 0)$, where $\overline{X}_n^* := \frac{1}{n}\sum_{i=1,\ldots,n} X_i^*$.

Suppose $\mu = 0$. Let $A_c := \{\liminf_{n\to\infty} n^{1/2}\overline{X}_n < -c\}$ for $0 < c < \infty$. By the law of the iterated logarithm, $P(A_c) = 1$. For $\omega \in A_c$, consider a subsequence $\{n_k : k \geq 1\}$ of $\{n : n \geq 1\}$ such that $n_k^{1/2}\overline{X}_{n_k}(\omega) \leq -c$ for all k. Then,

$$\begin{aligned} & n_k^{1/2}(\widehat{\mu}_{n_k}^* - \widehat{\mu}_{n_k}(\omega)) \\ = \;& \max(n_k^{1/2}(\overline{X}_{n_k}^* - \overline{X}_{n_k}(\omega)) + n_k^{1/2}\overline{X}_{n_k}(\omega), 0) - \max(n_k^{1/2}\overline{X}_{n_k}(\omega), 0) \\ \leq \;& \max(n_k^{1/2}(\overline{X}_{n_k}^* - \overline{X}_{n_k}(\omega)) - c, 0) \\ \xrightarrow{d} \;& \max(Z - c, 0) \text{ as } k \to \infty \text{ conditional on } \{\widehat{F}_n : n \geq 1\} \\ \leq \;& \max(Z, 0), \end{aligned} \quad (2)$$

where the last inequality is strict with positive probability and the convergence in distribution holds by a triangular array central limit theorem. So, along the subsequence $\{n_k\}$, $n_k^{1/2}(\widehat{\mu}_{n_k}^* - \widehat{\mu}_{n_k}(\omega)) \overset{d}{\nrightarrow} \max(Z,0)$ as $k \to \infty$ conditional on $\{\widehat{F}_{n_k} : k \geq 1\}$. Hence, $n^{1/2}(\widehat{\mu}_n^* - \widehat{\mu}_n(\omega)) \overset{d}{\nrightarrow} \max(Z,0)$ as $n \to \infty$ conditional on $\{\widehat{F}_n : n \geq 1\}$. This is true for all $\omega \in A_c$. We conclude that with probability one (with respect to the randomness in $\{\widehat{F}_n : n \geq 1\}$), the bootstrap distribution is not consistent.

Note that the bootstrap also is not correct when $\mu = 0$ for sample paths $\omega \in B_c := \{\limsup_{n\to\infty} n^{1/2}\overline{X}_n > c\}$ for any $0 < c < \infty$ and sample sizes $\{n_m : m \geq 1\}$ for which $n_m^{1/2}\overline{X}_{n_m}(\omega) \geq c$ for all m. In this case, we have

$$\begin{aligned} n_m^{1/2}(\widehat{\mu}_{n_m}^* - \widehat{\mu}_{n_m}(\omega)) &= \max(n_m^{1/2}(\overline{X}_{n_m}^* - \overline{X}_{n_m}(\omega)), -n_m^{1/2}\overline{X}_{n_m}(\omega)) \\ &\leq \max(n_m^{1/2}(\overline{X}_{n_m}^* - \overline{X}_{n_m}(\omega)), -c) \\ &\xrightarrow{d} \max(Z, -c) \text{ as } m \to \infty \text{ conditional on } \{\widehat{F}_n : n \geq 1\} \\ &\leq \max(Z, 0), \end{aligned} \quad (3)$$

where the last inequality is strict with positive probability. Note that $P(B_c) = 1$ for all $0 < c < \infty$. Thus, the bootstrap is incorrect both when $n^{1/2}\overline{X}_n(\omega)$ is negative for n large and when $n^{1/2}\overline{X}_n(\omega)$ is positive for n large. In both cases, the bootstrap distribution is too small (i.e., has too much mass to the left) when $\mu = 0$.

One can see why the bootstrap fails when $\mu = 0$ by inspecting equations (2) and (3) and utilizing the fact that $n^{1/2}(\overline{X}_n^* - \overline{X}_n(\omega))$ and $n^{1/2}(\overline{X}_n - \mu)$ have the same $N(0,1)$ distribution asymptotically. When $\overline{X}_n(\omega) = -c$ for $c > 0$, then $n^{1/2}(\widehat{\mu}_n^* - \widehat{\mu}_n(\omega)) = 0$ whenever $\overline{X}_n^* - \overline{X}_n(\omega) < c$, whereas $n^{1/2}(\widehat{\mu}_n - \mu) = 0$ whenever $\overline{X}_n - \mu < 0$.

Because $c > 0$, $n^{1/2}(\widehat{\mu}_n^* - \widehat{\mu}_n(\omega))$ has a higher probability of equalling zero than does $n^{1/2}(\widehat{\mu}_n - \mu)$. Alternatively, when $\overline{X}_n(\omega) = c$ for $c > 0$, then $n^{1/2}(\widehat{\mu}_n^* - \widehat{\mu}_n(\omega)) = \max(n^{1/2}(\overline{X}_n^* - \overline{X}_n(\omega)), -c)$, whereas $n^{1/2}(\widehat{\mu}_n - \mu) = \max(n^{1/2}(\overline{X}_n - \mu), 0)$. Because $-c < 0$, the distribution of $n^{1/2}(\widehat{\mu}_n^* - \widehat{\mu}_n(\omega))$ is to the left of that of $n^{1/2}(\widehat{\mu}_n - \mu)$.

This counterexample to the bootstrap generalizes to a wide variety of estimation problems that have considerable relevance in applications. For example, in models with random coefficients, it is often the case that the estimated variances of some of the random coefficients are small and, hence, the true variances of some of these coefficients may be zero. If any of the coefficient variances are zero, then the true parameter is on the boundary of the parameter space and the bootstrap is not consistent by an analogous argument to that given above. For brevity, we do not provide the details. See Andrews (1997) for some general results providing the asymptotic distribution of extremum estimators, including maximum likelihood estimators, minimum distance estimators, *etc.*, when the true parameter is on the boundary of the parameter space. These results cover random coefficient models. Such results are needed to demonstrate the inconsistency of the bootstrap in more general cases than the simple example provided above.

Bickel and Freedman (1981, Sec. 6) list three conditions for the bootstrap distribution of a statistic to be consistent in iid contexts. The first is weak convergence of the statistic when $X_i \sim G$ for all distributions G in a neighborhood of the true distribution F. The second is uniform weak convergence over distributions G in a neighborhood of the true distribution F. The third is continuity of the mapping from the underlying distribution G to the asymptotic distribution of the statistic. Bickel and Freedman provide two counterexamples to the bootstrap that violate the second condition, viz., uniformity. The counterexample given above violates the third condition, viz., continuity.

Bickel and Freedman's first counterexample to the bootstrap is a U-statistic of degree two in which the kernel $\omega(x, x)$ does not satisfy the condition $\int \omega^2(x, x) dF(x) < \infty$, where F denotes the true distribution of the data. Their second example is the largest order statistic from an iid sample of uniform $(0, \theta)$ random variables. This example is extended in Bickel, Götze, and van Zwet (1997, Example 3). Other counterexamples to the bootstrap include: extrema for unbounded distributions (see Athreya and Fukuchi (1994) and Deheuvels, Mason, and Shorack (1993)); the sample mean in the case of infinite variance random variables (see Babu (1984) and Artheya (1987)); Hodges' superefficient estimator (see Beran (1984)); degenerate U and V statistics (see Bretagnolle (1983)); nondifferentiable functions of the empirical distribution function (see Beran and Srivastava (1985) and Dümbgen (1993)); and the nonparametric kernel estimator of the mode of a smooth unimodal density when the smoothing parameter (for both the estimator and the bootstrap) is chosen to be optimal for the estimation problem (see Romano (1988)).

The counterexample to the bootstrap introduced above, based on a parameter being on the boundary of the parameter space, seems simpler and more relevant to many applications than most of the counterexamples just listed.

2 Alternatives to the Bootstrap

We now suggest three methods for obtaining consistent estimators of the asymptotic distribution of the normalized maximum likelihood estimator, $n^{1/2}(\widehat{\mu}_n - \mu)$, in the iid $N(\mu, 1)$ counterexample given above. These methods are designed to be consistent whether or not the true parameter is on the boundary. The methods generalize to the problem of an arbitrary extremum estimator when the true parameter may be on the boundary of the parameter space; see Andrews (1997).

The first method is as follows. Let $\{\eta_n : n \geq 1\}$ be a sequence of positive random variables (possibly constants) that satisfies

$$P(\lim_{n\to\infty} \eta_n = 0 \text{ and } \liminf_{n\to\infty} \eta_n (n/(2\ln\ln n))^{1/2} > 1) = 1. \tag{4}$$

If $\widehat{\mu}_n \leq \eta_n$, then we estimate the asymptotic distribution of $n^{1/2}(\widehat{\mu}_n - \mu)$ to be $\max(Z, 0)$. Otherwise, we estimate the asymptotic distribution to be Z. (Note that the η_n 's could be chosen to be the critical values for a sequence of one-sided tests of $H_0 : \mu = 0$ versus $H_1 : \mu > 0$ whose significance levels converge to zero as $n \to \infty$ at a rate such that (4) holds.)

This estimator of the asymptotic distribution is strongly consistent, because

$$\begin{aligned} & P\left(\limsup_{n\to\infty} (\widehat{\mu}_n - \eta_n) \leq 0\right) \\ = \ & P\left(\limsup_{n\to\infty} \left(\max((2n\ln\ln n)^{-1/2} \sum_{i=1}^{n} X_i, 0) - \eta_n (n/(2\ln\ln n))^{1/2}\right) \leq 0\right) \\ = \ & \begin{cases} 0 & \text{if } \mu > 0 \\ 1 & \text{if } \mu = 0 \end{cases} \end{aligned} \tag{5}$$

by the law of the iterated logarithm. Equation (5) also holds with the $\limsup_{n\to\infty}$ replaced by $\liminf_{n\to\infty}$.

This method of estimating the asymptotic distribution can be generalized to the case of an arbitrary extremum estimator with a parameter space that is defined by linear or nonlinear inequality constraints by specifying a criterion for each inequality constraint to assess whether it is binding or not. The method can be applied when the data are iid, as well as when the data exhibit temporal dependence, including stochastic and deterministic time trends. See Andrews (1997) for details.

The second method is a subsample method introduced by Wu (1990) and extended by Politis and Romano (1994) to cover cases where the statistic of interest has *some* asymptotic distribution, not necessarily normal, such as that which arises when the true parameter is on the boundary of the parameter space. Also see Bickel, Götze, and van Zwet (1997). The method is applicable in iid contexts, as well as in stationary time series contexts; see Politis and Romano (1994). A random subsampling variant of the procedure is also available; see Politis and Romano (1994, Sec. 2.2).

The third method is a variant of the bootstrap in which bootstrap samples of size n_0 ($< n$), rather than n, are employed. This method has been used previously as a means of fixing the bootstrap in the U-statistic counterexample of Bickel and

Freedman (1981) by Bretagnolle (1983) and in the sample mean with infinite variance random variables counterexample of Babu (1984); see Arcones (1990), who attributes the idea to an unpublished paper of Athreya. See Bickel, Götze, and van Zwet (1997) for further applications and analysis of this method.

The idea is to use the bootstrap distribution of $n_0^{1/2}(\widehat{\mu}_{n_0}^* - \widehat{\mu}_n)$ to estimate the distribution of $n^{1/2}(\widehat{\mu}_n - \mu)$, where $\widehat{\mu}_{n_0}^* := \max(\overline{X}_{n_0}^*, 0)$, $\overline{X}_{n_0}^* := \frac{1}{n_0}\sum_{i=1,\ldots,n_0} X_i^*$, and $\{X_i^* : i \leq n_0\}$ are iid with $X_i^* \sim \widehat{F}_n$. This variant of the bootstrap is consistent with probability one if $n_0 \to \infty$ and $n_0(\ln\ln n)/n \to 0$ as $n \to \infty$. The reason is that

$$\begin{aligned} & n_0^{1/2}(\widehat{\mu}_{n_0}^* - \widehat{\mu}_n) \\ = \; & \max(n_0^{1/2}(\overline{X}_{n_0}^* - \overline{X}_n) + n_0^{1/2}(\overline{X}_n - \mu), -n_0^{1/2}\mu) - n_0^{1/2}(\widehat{\mu}_n - \mu) \\ = \; & \max(n_0^{1/2}(\overline{X}_{n_0}^* - \overline{X}_n) + o(1), -n_0^{1/2}\mu) + o(1) \\ \overset{d}{\to} \; & \begin{cases} Z & \text{if } \mu > 0 \\ \max(Z, 0) & \text{if } \mu = 0 \end{cases} \text{ as } n \to \infty \text{ conditional on } \{\widehat{F}_n : n \geq 1\}, \end{aligned} \tag{6}$$

where the second equality holds with probability one by the law of the iterated logarithm and the convergence in distribution holds by the central limit theorem for triangular arrays of row-wise iid random variables.

3 Confidence Intervals

Here we consider the behavior of standard bootstrap confidence intervals for μ in the $N(\mu, 1)$ example discussed above. Because the variance is assumed to be known, there is no need to studentize the maximum likelihood estimator $\widehat{\mu}_n$ in this example. In any event, studentizing $\widehat{\mu}_n$ would not affect the first order asymptotic properties obtained here.

We consider two types of bootstrap *percentile* confidence intervals for μ, denoted CI_1 and CI_2. The first is based on percentiles of $n^{1/2}(\widehat{\mu}_n^* - \widehat{\mu}_n)$ and the second is based on percentiles of $\widehat{\mu}_n^*$. The first is what Hall (1992) refers to as the bootstrap percentile confidence interval and the second is what Efron and Tibshirani (1993) refer to as the bootstrap percentile confidence interval (and what Hall (1992) refers to as the "other" bootstrap percentile confidence interval).

The first bootstrap percentile confidence interval is

$$CI_1 = [\widehat{\mu}_n - \widehat{t}_{1\alpha_1}/n^{1/2}, \widehat{\mu}_n + \widehat{t}_{2\alpha_2}/n^{1/2}], \tag{7}$$

where $\widehat{t}_{1\alpha_1}$ is the $(1-\alpha_1)$-th quantile of $n^{1/2}(\widehat{\mu}_n^* - \widehat{\mu}_n)$ and $\widehat{t}_{2\alpha_2}$ is minus the α_2-th quantile of $n^{1/2}(\widehat{\mu}_n^* - \widehat{\mu}_n)$. We assume that $\alpha_1 + \alpha_2 = \alpha$ and that α_1, α_2, and α are in $[0, 1/2)$. An equal-tailed confidence interval is obtained by taking $\alpha_1 = \alpha_2$; a one-sided confidence interval is obtained by taking $\alpha_1 = 0$ or $\alpha_2 = 0$ (in which case $\widehat{t}_{1\alpha_1} := \infty$ or $\widehat{t}_{2\alpha_2} := \infty$ respectively); and a symmetric confidence interval is obtained by taking α_1 and α_2 such that $\widehat{t}_{1\alpha_1} = \widehat{t}_{2\alpha_2}$.

The coverage probability of CI_1 when $\mu = 0$ is

$$P(-\widehat{t}_{2\alpha_2} \leq n^{1/2}\widehat{\mu}_n \leq \widehat{t}_{1\alpha_1}). \tag{8}$$

To determine the limit of this probability as $n \to \infty$, we separately consider the cases where $\overline{X}_n \geq 0$ and $\overline{X}_n < 0$. All probabilities below refer to the case where $\mu = 0$.

Let $\widehat{z}_\alpha$ denote the αth quantile of $n^{1/2}(\overline{X}_n^* - \overline{X}_n)$ conditional on $\widehat{F}_n$. We have $\widehat{z}_\alpha \to z_\alpha$ as $n \to \infty$ with probability one, where z_α denotes the αth quantile of a standard normal distribution. Because α_1 and α_2 are each less than 1/2, $\widehat{z}_{1-\alpha_1} > 0$ and $\widehat{z}_{\alpha_2} < 0$ for n large with probability one.

When $\overline{X}_n \geq 0$, we have $\widehat{\mu}_n = \overline{X}_n \geq 0$, $n^{1/2}(\widehat{\mu}_n^* - \widehat{\mu}_n) = \max(n^{1/2}(\overline{X}_n^* - \overline{X}_n), -n^{1/2}\overline{X}_n)$, $\widehat{t}_{1\alpha_1} = \max(\widehat{z}_{1-\alpha_1}, -n^{1/2}\overline{X}_n)$, and $-\widehat{t}_{2\alpha_2} = \max(\widehat{z}_{\alpha_2}, -n^{1/2}\overline{X}_n)$. Thus, $-\widehat{t}_{2\alpha_2} \leq n^{1/2}\widehat{\mu}_n$ for n large with probability one because $\widehat{z}_{\alpha_2} \leq 0$ and $n^{1/2}\widehat{\mu}_n \leq \widehat{t}_{1\alpha_1}$ iff $n^{1/2}\overline{X}_n \leq \widehat{z}_{1-\alpha_1}$. When $\overline{X}_n < 0$, we have $\widehat{\mu}_n = 0$, $n^{1/2}(\widehat{\mu}_n^* - \widehat{\mu}_n) = \max(n^{1/2}\overline{X}_n^*, 0) \geq 0$, $\widehat{t}_{1\alpha_1} \geq 0$, and $-\widehat{t}_{2\alpha_2} = \max(\widehat{z}_{\alpha_2} + n^{1/2}\overline{X}_n, 0) = 0$, where the last equality holds for n large with probability one because $\widehat{z}_{\alpha_2} \leq 0$. Hence, $-\widehat{t}_{2\alpha_2} \leq \widehat{\mu}_n \leq \widehat{t}_{1\alpha_1}$ always holds when $\overline{X}_n < 0$.

We conclude that the coverage probability of CI_1 when $\mu = 0$ satisfies

$$P(-\widehat{t}_{2\alpha_2} \leq n^{1/2}\widehat{\mu}_n \leq \widehat{t}_{1\alpha_1}) = P(n^{1/2}\overline{X}_n > \max(\widehat{z}_{1-\alpha_1}, 0)) \to \alpha_1. \tag{9}$$

The confidence interval CI_1 never misses to the left. If $\alpha_2 > 0$, then its asymptotic coverage probability exceeds its nominal level $1 - \alpha$ when $\mu = 0$.

Next, we consider the second bootstrap percentile confidence interval. It is defined to be

$$CI_2 = [q_{\alpha_1}^*, q_{1-\alpha_2}^*], \tag{10}$$

where q_α^* is the αth quantile of the distribution of $\widehat{\mu}_n^*$ conditional on $\widehat{F}_n$. Because $\widehat{\mu}_n^* \geq 0$, we have $q_{1-\alpha_2}^* \geq 0$ and the confidence interval CI_1 never misses the true value $\mu = 0$ to the left. Hence, CI_2 covers the true value $\mu = 0$ unless $q_{\alpha_1}^* > 0$. We have $q_{\alpha_1}^* > 0$ iff $P(\widehat{\mu}_n^* \leq 0|\widehat{F}_n) < \alpha_1$ iff $P(\overline{X}_n^* \leq 0|\widehat{F}_n) < \alpha_1$ iff $P(n^{1/2}(\overline{X}_n^* - \overline{X}_n) \leq -n^{1/2}\overline{X}_n|\widehat{F}_n) < \alpha_1$ iff $-n^{1/2}\overline{X}_n < \widehat{z}_{\alpha_1}$, where $P(\cdot|\widehat{F}_n)$ denotes probability conditional on $\widehat{F}_n$. A central limit theorem gives $P(-n^{1/2}\overline{X}_n < \widehat{z}_{\alpha_1}) \to \alpha_1$. Thus, we conclude that the asymptotic coverage probability of CI_2 is α_1, the same as for CI_1. If $\alpha_2 > 0$, then the asymptotic coverage probability of CI_2 exceeds its nominal level $1 - \alpha$ when $\mu = 0$.

References

Andrews, D. W. K. (1997): "Estimation When a Parameter Is on a Boundary," unpublished manuscript, Cowles Foundation for Research in Economics, Yale University.

Arcones, M. (1990): *On the Asymptotic Theory of the Bootstrap*. Ph. D. dissertation. The City University of New York.

Artheya, K. B. (1987): "Bootstrap of the mean in the infinite variance case," *Annals of Statistics*, 15, 724–731.

Artheya, K. B. and J. Fukuchi (1994): "Bootstrapping Extremes of I.I.D. Random Variables," *Proceedings of the Conference on Extreme Value Theory*. NIST.

Babu, J. (1984): "Bootstrapping Statistics with Linear Combinations of Chi-squares as Weak Limit," *Sankhya,* 46, 86-93.

Beran, R. (1984): "Bootstrap Methods in Statistics," *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 86, 14-30.

Beran, R. and M. S. Srivastava (1985): "Bootstrap Tests and Confidence Regions for Functions of a Covariance Matrix," *Annals of Statistics*, 13, 95–115.

Bickel, P. J. and D. Freedman (1981): "Some Asymptotic Theory for the Bootstrap," *Annals of Statistics,* 9, 1196-1217.

Bickel, P. J., F. Götze, and W. R. van Zwet (1997): "Resampling Fewer than n Observations: Gains, Losses, and Remedies for Losses," *Statistica Sinica*, 7, 1–31.

Bretagnolle, J. (1983): "Lois Limites du Bootstrap de Certaines Fonctionnelles," *Annales de L'institut Henri Poincaré*, Sec. B, 19, 281-296.

Deheuvels, P., D. Mason, and G. Shorack (1993): "Some Results on the Influence of Extremes on the Bootstrap," *Annales de L'institut Henri Poincaré*, 29, 83–103.

Dümbgen, L. (1993): "On Nondifferentiable Functions and the Bootstrap," *Probability Theory and Related Fields*, 95, 125–140.

Efron, B. (1979): "Bootstrap Methods: Another Look at the Jackknife," *Annals of Statistics,* 7, 1-26.

Efron, B. and R. Tibshirani (1993): *An Introduction to the Bootstrap*. New York: Chapman and Hall.

Freedman, D. (1981): "Bootstrapping Regression Models," *Annals of Statistics*, 9, 1218-1228.

Hall, P. (1992): *The Bootstrap and Edgeworth Expansion.* New York: Springer–Verlag.

Politis, D. N. and J. P. Romano (1994): "Large Sample Confidence Regions Based on Subsamples under Minimal Assumptions," *Annals of Statistics,* 22, 2031-2050.

Romano, J. (1988): "Bootstrapping the Mode," *Annals of the Institute of Statistical Mathematics,* 40, 565-586.

Wu, C. F. J. (1990): "On the Asymptotic Properties of the Jackknife Histogram," *Annals of Statistics,* 18, 1438-1452.