Ann. Inst. Statist. Math.
Vol. 48, No. 4, 663–673 (1996)

BAYESIAN NONPARAMETRIC PREDICTIVE INFERENCE AND BOOTSTRAP TECHNIQUES

P. MULIERE AND P. SECCHI

Università di Pavia, Dipartimento di Economia Politica e Metodi Quantitativi, Via San Felice, 5, I-27100 Pavia, Italy

(Received January 23, 1995; revised January 5, 1996)

Abstract. We address the question as to whether a prior distribution on the space of distribution functions exists which generates the posterior produced by Efron's and Rubin's bootstrap techniques, emphasizing the connections with the Dirichlet process. We also introduce a new resampling plan which has two advantages: prior opinions are taken into account and the predictive distribution of the future observations is not forced to be concentrated on observed values.

Key words and phrases: Bootstrap techniques, Dirichlet process, nonparametric predictive inference.

1. Introduction

The bootstrap resampling plan introduced by Efron (1979) has a Bayesian counterpart called by Rubin (1981) the Bayesian bootstrap. Both resampling plans are asymptotically equivalent (Lo (1987), Weng (1989)) and first order equivalent from the predictive point of view. We investigate the question as to whether the posterior distributions obtained by means of the bootstrap procedures arise via Bayes Theorem from a prior on the space of distribution functions. The fact that the Bayesian bootstrap "gives zero probability to the event that a future observation is unequal to the observed values in the sample" (Meeden (1993)) led some Bayesian authors to question its applicability and to suggest modifications to the basic procedure. We also suggest a new generalization of the Bayesian bootstrap which takes into account prior opinions and has moreover the appealing property that the predictive distribution for a future observation is not necessarily concentrated on the observed values.

The paper is organized as follows. In the next section we introduce the bootstrap resampling techniques of Efron and Rubin in a general Bayesian nonparametric context where one wants to approximate the posterior distribution of a statistical functional $\phi(F)$ with F a random distribution function. We characterize the prior distributions for F which generate Efron's and Rubin's bootstraps, emphasizing the connections with the Dirichlet process. In Section 3 we propose a new bootstrap technique. A couple of applications are discussed in the last section of the paper.

Let us set some notation and terminology. The definition of a distribution Beta(α, β) requires for both parameters to be strictly positive. If $\alpha = 0$ and $\beta > 0$, we indicate by Beta$(0, \beta)$ the distribution function of the point mass at 0, whereas if $\alpha > 0$ and $\beta = 0$, we define Beta$(\alpha, 0)$ to be the distribution function of the point mass at 1. Analogously, for every integer $n > 0$, Binomial$(n, 0)$ will indicate the distribution function of the point mass at 0.

2. Bootstrap and prior distribution for F

Let $\{X_n\}$ be an exchangeable sequence of real random variables (r.v.) defined on a probability space $(\Omega, \mathcal{F}, P)$. De Finetti's Representation Theorem guarantees the existence of a random distribution function F conditionally on which the variables of the sequence $\{X_n\}$ are independent and identically distributed (i.i.d.) with distribution F.

In the Bayesian context the bootstrap procedures provide methods for approximating the conditional distribution

$$\mathcal{L}(\phi(F, \boldsymbol{X}) \mid X_1, \ldots, X_n) \tag{2.1}$$

where, for clarity of exposition, we indicated with $\boldsymbol{X}$ the sample $X_1, \ldots, X_n$ and $\phi(F, \boldsymbol{X})$ is a functional depending on both F and $\boldsymbol{X}$.

Efron's bootstrap suggests to approximate the conditional distribution (2.1) by means of

$$\mathcal{L}(\phi(F_n^*, \boldsymbol{X}) \mid X_1, \ldots, X_n) \tag{2.2}$$

where F_n^* is the empirical distribution of an i.i.d. sample $X_1^*, \ldots, X_n^*$ from the empirical distribution function F_n of $X_1, \ldots, X_n$. In particular, for every Borel set B, the conditional distribution

$$\mathcal{L}(F(B) \mid X_1, \ldots, X_n) \tag{2.3}$$

is approximated by

$$\mathcal{L}(F_n^*(B) \mid X_1, \ldots, X_n) = \frac{1}{n}\text{Binomial}(n, F_n(B)) \tag{2.4}$$

where n^{-1}Binomial$(n, F_n(B))$ indicates the distribution of a r.v. $Y \in [0, 1]$ such that nY has Binomial$(n, F_n(B))$ distribution.

The Bayesian bootstrap suggests instead to approximate (2.1) by means of the conditional distribution

$$\mathcal{L}(\phi(F_n^R, \boldsymbol{X}) \mid X_1, \ldots, X_n)$$

where, given $X_1, \dots, X_n$, F_n^R is the random distribution function defined by setting, for every $x \in \Re$,

$$F_n^R(x) = \frac{1}{\sum_{i=1}^n V_i} \sum_{i=1}^n V_i I_{[X_i,\infty)}(x)$$

and $\{V_n\}$ is a sequence of i.i.d. random variables with exponential distribution of parameter 1 independent of $\{X_n\}$. Therefore, for every Borel set B, the conditional distribution of $F(B)$, given $X_1, \dots, X_n$, is approximated in this case by

$$\mathcal{L}(F_n^R(B) \mid X_1, \dots, X_n) = \mathrm{Beta}(nF_n(B), n[1 - F_n(B)]). \tag{2.5}$$

Since, for every Borel set B and for all n,

$$P[X_{n+1} \in B \mid X_1, \dots, X_n] = E[F(B) \mid X_1, \dots, X_n], \tag{2.6}$$

approximations (2.4) and (2.5) imply that both Efron's and Rubin's bootstraps evaluate the conditional probability that the next observation falls in B equal to the frequency according to which the past observations fell in B. In this sense they are first order equivalent from the predictive point of view.

The Bayesian approach to the evaluation of the conditional probability (2.6) requires to elicit a prior distribution for F on the space of distribution functions and then to use the posterior distribution of F for the computation of the expected value appearing in (2.6). An interesting prior for F was introduced by Ferguson (1973) in a fundamental paper on a Bayesian approach to nonparametric statistics. We will indicate this prior, called Dirichlet process, by $\mathcal{D}(kF_0)$ where $k > 0$ is a real number and F_0 is a proper distribution function. F_0 can be interpreted as the prior guess at F whereas k is the 'measure of faith' in this guess. For the definition of the Dirichlet process and a review of its salient features we refer to the seminal papers of Ferguson (1973, 1974).

In the rest of this section we want to investigate the question as to when the bootstrap procedures are in agreement with the Bayesian approach. The next theorem characterizes the priors for F such that, for every Borel set B and every n, the posterior distribution of $F(B)$, conditionally on $X_1, \dots, X_n$, is given by (2.4) or by (2.5).

THEOREM 2.1. *Given a random distribution function F, let $\{X_n\}$ be a sequence of r.v.'s conditionally i.i.d. with distribution F. For every Borel set B and for every n,*

$$\mathcal{L}(F(B) \mid X_1, \dots, X_n) \tag{2.7}$$

is equal to (2.4) *or* (2.5) *if and only if F concentrates on a random point.*

PROOF.

Sufficiency. Assume that, given a r.v. X, F concentrates on X. Then, for all $n > 0$,

$$P[X = X_1 = \cdots = X_n] = 1.$$

Fix a Borel set B and note that, for all n,

$$\begin{aligned}\mathcal{L}(F(B) \mid X_1, \ldots, X_n) &= \begin{cases} I_{[1,\infty)}(\cdot) & \text{if } X_1 \in B \\ I_{[0,\infty)}(\cdot) & \text{if } X_1 \notin B \end{cases} \\ &= \text{Beta}(nF_n(B), n[1 - F_n(B)]) \\ &= \frac{1}{n}\text{Binomial}(n, F_n(B))\end{aligned}$$

since $P(F_1(B) = F_n(B)) = 1$.

Necessity. Assume that (2.4) or (2.5) holds for every Borel set B and for every n. In particular, for every given Borel set B,

$$\mathcal{L}(F(B) \mid X_1) = \begin{cases} I_{[1,\infty)}(\cdot) & \text{if } X_1 \in B \\ I_{[0,\infty)}(\cdot) & \text{if } X_1 \notin B. \end{cases}$$

Therefore, given X_1, F concentrates on X_1. □

Remark 2.1. The previous characterization theorem is somewhat related to a result due to Regazzini (1978) and Lo (1991) which states that, for every $n \geq 1$ and for every Borel set B,

$$P[X_{n+1} \in B \mid X_1, \ldots, X_n] = \frac{k}{k+n} F_0(B) + \frac{n}{k+n} F_n(B)$$

with $k > 0$ and F_0 a distribution function, if and only if F is a Dirichlet process $\mathcal{D}(kF_0)$. In fact Theorem 2.1, could be considered as an extension of this result to the case $k = 0$.

A random distribution F concentrated on a random point is of no interest for the statistician since it implies that the exchangeable random variables of the sequence $\{X_n\}$ are all equal with probability one. However an F of this type can be regarded as a limit of a Dirichlet process $\mathcal{D}(kF_0)$ when $k \to 0$ (Ferguson (1974)). Therefore it seems that bootstrap procedures are justifiable in the Bayesian context when the weight k given to the prior guess F_0 at F is extremely small. Lo (1987) has shown that the approximations (2.4) and (2.5) for the posterior distribution of $F(B)$, given $X_1, \ldots, X_n$, are also reasonable when n is large and the prior for F is a Dirichlet process $\mathcal{D}(kF_0)$. Even in this case the weight given to the prior opinion elicited with F_0 becomes negligible. In the next section we want to suggest a different bootstrap procedure which takes into account the prior opinion F_0 and can be considered as an extension of Efron's and Rubin's bootstraps.

3. A new bootstrap technique

In this section we assume that the random distribution function F conditionally on which the r.v.'s of the sequence $\{X_n\}$ are i.i.d. is a Dirichlet process

$\mathcal{D}(kF_0)$, with $k > 0$ and F_0 a proper distribution function. We want to suggest a resampling procedure with the aim of approximating the conditional distribution

$$\mathcal{L}(\phi(F) \mid X_1, \ldots, X_n) \tag{3.1}$$

where $\phi(F)$ is a functional depending on F. In particular we will consider two types of functional:

Q$_1$: $\phi_q(F) = \inf\{t \in \Re : F(t) \geq q\}$ where $0 < q < 1$;
Q$_2$: $\phi_h(F) = \int h dF$.

In principle the conditional distribution function (3.1) can be computed by means of Bayes Theorem. For functionals of type Q$_2$ this has been studied by Hannum *et al.* (1981) and by Cifarelli and Regazzini (1990). Their results, however, are not easy to handle analytically. Therefore arises the need for an approximating technique.

Our proposal stems from the fact that it is trivial to simulate a Dirichlet process when the parameter F_0 is discrete with finite support as it is made clear by the following lemma which we state without proof.

LEMMA 3.1. *Let F_0 be a discrete distribution function with support $\{z_1, \ldots, z_r\}$ in $\Re$. For $i = 1, \ldots, r$, let p_i be the probability which F_0 assigns to z_i. Assume that $V_1, \ldots, V_r$ are r independent r.v.'s such that, for $i = 1, \ldots, r$,*

$$\mathcal{L}(V_i) = \text{Gamma}[kp_i, 1]$$

where $k > 0$. If F is the random distribution function defined, for every $x \in \Re$, by setting

$$F(x) = \frac{1}{\sum_{i=1}^r V_i} \sum_{i=1}^r V_i I_{[z_i, \infty)}(x),$$

then F is a Dirichlet process $\mathcal{D}(kF_0)$.

Given a sample $X_1, \ldots, X_n$ from a Dirichlet process F with parameter kF_0, the posterior distribution of F is again a Dirichlet process with parameter $kF_0 + nF_n$. If F_0 is a discrete distribution with finite support, the parameter $(k+n)^{-1}(kF_0 + nF_n)$ is also discrete; let $\{z_1, \ldots, z_r\}$ be the finite support of this last distribution with corresponding probability masses $\{p_1, \ldots, p_r\}$. Then

$$\mathcal{L}(\phi(F) \mid X_1, \ldots, X_n) = \mathcal{L}\left(\phi\left(\frac{1}{\sum_{i=1}^r V_i} \sum_{i=1}^r V_i I_{[z_i, \infty)}\right) \mid X_1, \ldots, X_n\right)$$

where, given $X_1, \ldots, X_n$, the r.v.'s $V_1, \ldots, V_r$ are independent and such that, for $i = 1, \ldots, r$,

$$\mathcal{L}(V_i \mid X_1, \ldots, X_n) = \text{Gamma}[(k+n)p_i, 1].$$

In this case it is immediately evident how to apply a Monte Carlo method in order to find an approximation of (3.1). However, in most situations of statistical interest, F_0 is not discrete so that the direct approach just described will not be applicable. When this happens, a possible way out is first to approximate the parameter $kF_0 + nF_n$ with a suitable bounded, monotone increasing, right continuous step function α^* such that $\alpha^*(-\infty) = 0$, and then to use the process $\mathcal{D}(\alpha^*)$ as an approximation of the posterior process $\mathcal{D}(kF_0 + nF_n)$. Rubin's bootstrap originates from the same idea by setting $\alpha^* = nF_n$. Our alternative proposal is to approximate $kF_0 + nF_n$ by $(n+k)F_m^*$ where F_m^* is the empirical distribution function generated by an i.i.d. sample of size m from $(n+k)^{-1}(kF_0 + nF_n)$.

In order to justify our proposal, let G be a Dirichlet process $\mathcal{D}(wG_0)$; when a sample $X_1, \ldots, X_n$ from a process $\mathcal{D}(kF_0)$ has been observed, set

$$w = n + k \quad \text{and} \quad G_0 = \frac{k}{n+k} F_0 + \frac{n}{n+k} F_n.$$

For any given m, let $\boldsymbol{X}^* = (X_1^*, \ldots, X_m^*)$ be an i.i.d. sample from the distribution G_0. Define G_m^* to be a random distribution which, conditionally on the empirical distribution F_m^* of $\boldsymbol{X}^*$, is a Dirichlet process $\mathcal{D}(wF_m^*)$. In agreement with the definition of Antoniak (1974), G_m^* is a mixture of Dirichlet processes; note that, for any given measurable partion $B_1, \ldots, B_r$ of $\Re$, the marginal distribution of $(G_m^*(B_1), \ldots, G_m^*(B_r))$ is a mixture of Dirichlet distributions with normalized Multinomial weights.

When m grows to infinity, the law of G_m^* weakly converges to the law of G. With the next two theorems we show that the distribution of $\phi(G_m^*)$ weakly converges to the distribution of $\phi(G)$ when ϕ is a functional of type Q_1 or Q_2.

THEOREM 3.1. *Let $0 < q < 1$ and G be a Dirichlet process $\mathcal{D}(wG_0)$. Then the distribution of $\phi_q(G_m^*)$ converges weakly to the distribution of $\phi_q(G)$, when $m \to \infty$.*

PROOF. We need to show that, if t is a continuity point of the distribution of $\phi_q(G)$,

$$\lim_{m\to\infty} P[\phi_q(G_m^*) \le t] = P[\phi_q(G) \le t]. \tag{3.2}$$

By Glivenko-Cantelli Lemma, the sequence $\{F_m^*\}$ converges uniformly to G_0 on a set A of probability one. Fix a continuity point t of G_0 and note that, conditionally on F_m^*, the random variable $G_m^*(t)$ has distribution

$$\mathrm{Beta}(wF_m^*(t), w(1 - F_m^*(t)))$$

which weakly converges to a $\mathrm{Beta}(wG_0(t), w(1-G_0(t)))$ on A, as $m \to \infty$. Therefore,

$$\begin{aligned}\lim_{m\to\infty} P[\phi_q(G_m^*) \le t \mid F_m^*] &= \lim_{m\to\infty} P[G_m^*(t) \ge q \mid F_m^*] \\ &= P[G(t) \ge q] = P[\phi_q(G) \le t]\end{aligned}$$

on a set of probability one. This implies that (3.2) holds when t is a continuity point for G_0. The proof is completed by showing that any discontinuity point of G_0 is also a discontinuity point of the distribution of $\phi_q(G)$. □

THEOREM 3.2. *Let G be a Dirichlet process $\mathcal{D}(wG_0)$ and h be a real valued, bounded and G_0-continuous function defined on $\Re$. Then $\phi_h(G_m^*)$ converges weakly to $\phi_h(G)$, when $m \to \infty$.*

PROOF. Apply Glivenko-Cantelli Lemma and Corollary 2.7 of Hannum *et al.* (1981). □

Assume that $X_1, \ldots, X_n$ is a sample from a Dirichlet process F with parameter kF_0 and $\phi(F)$ is a functional of type Q_1 or Q_2. The previous results imply that the conditional distribution of $\phi(F)$, given $X_1, \ldots, X_n$, can be approximated, for m large, by the conditional distribution

$$\mathcal{L}(\phi(G_m^*) \mid X_1, \ldots, X_n)$$

where, given $X_1, \ldots, X_n$ and the empirical distribution function F_m^* of an i.i.d. sample $X_1^*, \ldots, X_m^*$ from the distribution

$$\frac{k}{n+k} F_0 + \frac{n}{n+k} F_n,$$

the process G_m^* is Dirichlet with parameter $(n+k)F_m^*$. Details of a resampling procedure supported by this argument will be introduced in the next section along with a couple of numerical examples. Note that the bootstrap technique suggested above approximates the conditional probability $P[X_{n+1} \in B \mid X_1, \ldots, X_n]$ by means of

$$E[G_m^*(B) \mid X_1, \ldots, X_n] = \frac{kF_0(B) + nF_n(B)}{k+n}.$$

This is the same predictive probability obtained by computing, via Bayes Theorem, the posterior distribution of F. Moreover, by taking into account the prior opinion elicited by the distribution F_0 with weight $k > 0$, this resampling plan does not force the future observation to be equal to one of the observed values as is the case with Efron's and Rubin's bootstraps. The technique has also the advantage of being fully consistent with the Bayesian approach since it can be considered as only a tool for approximating numerically a 'true' posterior distribution when this is analytically hard to manage.

4. Numerical illustrations

We now want to describe a resampling plan which has the aim of computing an approximation for (3.1) and is supported by the arguments of the previous section. The procedure will be tested with two applications.

Assume that a sample $X_1 = x_1, \ldots, X_n = x_n$ has been observed from a random distribution F. Elicit the prior opinion about F by a proper distribution function F_0, the prior 'guess' at F, and by a positive number k, the 'measure of faith' in this guess. In order to build a distribution function which approximates $\mathcal{L}(\phi(F) \mid X_1 = x_1, \ldots, X_n = x_n)$ we propose to follow these steps:

1. Generate m observations $x_1^*, \ldots, x_m^*$ from $(n+k)^{-1}(kF_0 + nF_n)$.
2. For $i = 1, \ldots, m$, generate v_i from a
$$\text{Gamma}\left(\frac{n+k}{m}, 1\right).$$
3. Compute the quantity
$$t = \phi\left(\frac{1}{\sum_{i=1}^m v_i} \sum_{i=1}^m v_i I_{[x_i^*, \infty)}(\cdot)\right).$$
4. Repeat steps (1), (2), (3) s times obtaining the quantities $t_1, \ldots, t_s$.
5. Approximate the conditional distribution function $\mathcal{L}(\phi(F) \mid X_1 = x_1, \ldots, X_n = x_n)$ by means of the empirical distribution function generated by $t_1, \ldots, t_s$.

Example 4.1. We observe $x_1 = 0.1$, $x_2 = 0.05$. Assuming that this is a sample from a random distribution F, we want to compute

$$\mathcal{L}\left(\int x dF(x) \mid x_1, x_2\right). \tag{4.1}$$

Our prior guess F_0 at F is a Uniform distribution on $[0, 1]$. To this guess we assign weights $k = 0$, 1, 2, 100. The procedure described above was then applied with $m = 300$ and $s = 5000$. For different values of k, the distributions approximating (4.1) are summarized in Table 1 by their mean, median, 75th and 95th quantile here indicated with q_{75} and q_{95} respectively.

For $k = 0$ our procedure is equivalent to Rubin's bootstrap. However, if the posterior distribution of F, given $X_1 = x_1$, $X_2 = x_2$, is a Dirichlet process $\mathcal{D}(2F_2)$, then one can verify that (4.1) is a Uniform distribution on $[0.05, 0.1]$ so that the values for the mean, the median and the quantiles can be computed analytically. On the other hand, when the prior distribution of F is a Dirichlet process $\mathcal{D}(kF_0)$, it's always possible to compute the mean of (4.1) (Ferguson (1973)). All these analytical results are reported in Table 1 between square brackets.

For comparison purposes we approximated the distribution (4.1) by means of a different technique. We assumed that the prior distribution for F is a Dirichlet process $\mathcal{D}(kF_0)$ and observed a sample of size 300 from a Pólya urn of parameter $kF_0 + 2F_2$ (Blackwell and MacQueen (1973), Lo (1988)). The mean of this sample can be viewed as a realization of a random variable having distribution (4.1). The procedure was then iterated 5000 times and the empirical distribution of the sample means thus obtained was considered as an approximation of (4.1). Results relative to this simulation are reported in Table 1 between round brackets.

The results computed by means of these different techniques look all very similar and they all confirm the obvious fact that the more k increases, the more relevant becomes the prior opinion elicited with F_0.

Table 1. Results of the experiments described in Example 4.1.

	Mean			Median		
$k = 0$	[0.0750]	0.0747	(0.0749)	[0.0750]	0.0745	(0.0750)
$k = 1$	[0.2166]	0.2156	(0.2152)	$\star$	0.1789	(0.1786)
$k = 2$	[0.2875]	0.2888	(0.2862)	$\star$	0.2675	(0.2654)
$k = 100$	[0.4916]	0.4918	(0.4891)	$\star$	0.4918	(0.4893)

	q_{75}			q_{95}		
$k = 0$	[0.0875]	0.0871	(0.0872)	[0.0975]	0.0975	(0.0973)
$k = 1$	$\star$	0.2813	(0.2815)	$\star$	0.4770	(0.4804)
$k = 2$	$\star$	0.3688	(0.3667)	$\star$	0.5462	(0.5279)
$k = 100$	$\star$	0.5156	(0.5114)	$\star$	0.5451	(0.5431)

Table 2. Exact values for the quantiles of a Gamma[20, 1].

	$q = 0.25$	$q = 0.50$	$q = 0.75$
$F =$ Gamma[20, 1]	16.83	19.67	22.81

Example 4.2. For the purpose of comparing our resampling plan with other bootstrap techniques recently introduced in the Bayesian literature we repeated an experiment originally due to Meeden (1993). Given a sample $X_1, \ldots, X_n$ from a random distribution function F, we want to estimate the 25th, the 50th and the 75th quantile of F. In the following experiments the sample $X_1, \ldots, X_n$ was generated by a Gamma[20, 1]. For $q = 0.25$, 0.5, 0.75, define, as before, the functional

$$\phi_q(F) = \inf\{t \in \Re : F(t) \geq q\}.$$

The exact values of $\phi_q(F)$ when $F =$ Gamma[20, 1] are reported in Table 2.

Two experiments, differing by the sample size, were performed; in the first one the sample size n was set equal to 11, in the second it was set to be 25. We first obtained an approximation of

$$\mathcal{L}(\phi_q(F) \mid X_1, \ldots, X_n) \tag{4.2}$$

by applying the procedure described at the beginning of the section with $m = 100$ and $s = 300$. We then estimated the 25th, the 50th and the 75th quantile of F by the mean of the corresponding distribution (4.2). As prior guess at F we considered four different distributions: Uniform[0, 60], Uniform[8.5, 30], LogNormal[2.97, 0.22] and Gamma[20, 1]. To each distribution we assigned three weights: $k = 0$, 1, 5. Each experiment was repeated 100 times; the average values of the estimators, indicated respectively by $\bar{q}_{25}$, $\bar{q}_{50}$ and $\bar{q}_{75}$, are reported in Table 3.

Note again that for $k = 0$ our procedure is equivalent to Rubin's Bayesian bootstrap. Since the values relative to this case were already computed by Meeden, for comparison purposes we reported them in Table 3.

Table 3. Results of the experiments described in Example 4.2.

		Prior Distribution			
		$U[0,60]$	$U[8.5,30]$	LogNormal$[2.97, 0.22]$	Gamma$[20,1]$
		$(n=11, n=25)$	$(n=11, n=25)$	$(n=11, n=25)$	$(n=11, n=25)$
	$k=0$	$(17.21, 16.96)$	$(17.21, 16.96)$	$(17.21, 16.96)$	$(17.21, 16.96)$
$\bar{q}_{25}$	$k=1$	$(17.19, 16.93)$	$(17.04, 16.82)$	$(16.99, 16.74)$	$(17.45, 16.97)$
	$k=5$	$(16.96, 16.87)$	$(16.44, 16.42)$	$(16.90, 16.81)$	$(17.07, 16.92)$
	$k=0$	$(19.89, 19.74)$	$(19.89, 19.74)$	$(19.89, 19.74)$	$(19.89, 19.74)$
$\bar{q}_{50}$	$k=1$	$(20.10, 19.78)$	$(19.77, 19.58)$	$(19.73, 19.48)$	$(19.96, 19.80)$
	$k=5$	$(20.84, 20.21)$	$(19.68, 19.42)$	$(19.63, 19.55)$	$(19.76, 19.68)$
	$k=0$	$(22.68, 22.83)$	$(22.68, 22.83)$	$(22.68, 22.83)$	$(22.68, 22.83)$
$\bar{q}_{75}$	$k=1$	$(23.46, 23.00)$	$(22.81, 22.71)$	$(22.85, 22.71)$	$(22.77, 22.94)$
	$k=5$	$(27.14, 24.21)$	$(23.17, 22.74)$	$(22.58, 22.69)$	$(22.72, 22.87)$

With this experiment Meeden (1993) compared Rubin's bootstrap with the smooth Bayesian bootstrap of Banks (1988) and his own Bayesian bootstrap based on a grid reaching the conclusion that all these procedures performed similarly.

Table 3 shows that the results obtained with our resampling technique are quite similar to those of Meeden especially when the weight given to the prior opinion is small compared to the sample size. In fact, even for $k = 5$ the average values of the estimators are close to those found with $k = 0$ with the possible exception of those relative to 75th quantile when the prior is Uniform$[0, 60]$.

Acknowledgements

We are grateful to an anonymous referee whose helpful comments contributed to improve the paper. The first version of this work was completed while both authors were visiting the School of Statistics and the Institute for Mathemathics and its Applications at the University of Minnesota. The research of Pietro Muliere was supported by C.N.R.

REFERENCES

Antoniak, C. (1974). Mixture of Dirichlet processes with application to bayesian nonparametric problems, *Ann. Statist.*, **12**, 636–645.

Banks, D. L. (1988). Histopline smoothing the Bayesian bootstrap, *Biometrika*, **75**, 673–684.

Blackwell, D. and MacQueen, J. B. (1973). Ferguson distributions via Pólya urn schemes, *Ann. Statist.*, **1**(2), 353–355.

Cifarelli, D. M. and Regazzini, E. (1990). Distribution functions of means of a Dirichlet process, *Ann. Statist.*, **18**(1), 429–442.

Efron, B. (1979). Bootstrap methods: another look at the Jackknife, *Ann. Statist.*, **7**(1), 1–26.

Ferguson, T. S. (1973). A Bayesian analysis of some nonparametric problems, *Ann. Statist.*, **1**(2), 209–230.
Ferguson, T. S. (1974). Prior distributions on spaces of probability measures, *Ann. Statist.*, **2**(4), 615–629.
Hannum, R. C., Hollander, M. and Langberg, N. A. (1981). Distributional results for random functionals of a Dirichlet process, *Ann. Probab.*, **9**(4), 665–670.
Lo, A. Y. (1987). A large sample study of the Bayesian bootstrap, *Ann. Statist.*, **15**(1), 360–375.
Lo, A. Y. (1988). A Bayesian bootstrap for a finite population, *Ann. Statist.*, **16**(4), 1684–1685.
Lo, A. Y. (1991). A characterization of the Dirichlet process, *Statist. Probab. Lett.*, **12**(3), 185–187.
Meeden, G. (1993). Noninformative nonparametric Bayesian estimation of quantiles, *Statist. Probab. Lett.*, **16**, 103–109.
Regazzini, E. (1978). Intorno ad alcune questioni relative alla definizione del premio secondo la teoria della credibilità, *Giornale dell'Istituto Italiano degli Attuari*, 77–89.
Rubin, D. B. (1981). The Bayesian bootstrap, *Ann. Statist.*, **9**(1), 130–134.
Weng, C. S. (1989). On a second order asymptotic property of the Bayesian bootstrap mean, *Ann. Statist.*, **17**(2), 705–710.