An urn-based Bayesian block bootstrap

Pasquale Cirillo · Pietro Muliere

Abstract   Block bootstrap has been introduced in the literature for resampling dependent data, i.e. stationary processes. One of the main assumptions in block boot-strapping is that the blocks of observations are exchangeable, i.e. their joint distribution is immune to permutations. In this paper we propose a new Bayesian approach to block bootstrapping, starting from the construction of exchangeable blocks. Our sampling mechanism is based on a particular class of reinforced urn processes.

Keywords   Block bootstrap · Exchangeability · Reinforced urn processes

1 Introduction

Block bootstrap (BB) has been introduced in Carlstein (1986) and Künsch (1989) to resample dependent data, for which standard bootstrapping rules are not correctly applicable. In particular it has been shown that this technique works well for a wide class of stationary processes without any parametric model assumptions, see Lahiri (2003).

The idea behind BB is simple: instead of resampling a single observation at a time, as in Efron's iid bootstrap (Efron 1979), one resamples blocks of (consecutive) observations. In this way, the dependence structure of the original observations is preserved within each block. The length of blocks is assumed to increase with sample size. As a consequence of this, if the data are generated by a weakly dependent process, the

---

P. Cirillo (✉)
Department of Statistics (IMSV), University of Bern, Sidlerstrasse 5, 3012 Bern, Switzerland
e-mail: pasquale.cirillo@stat.unibe.ch

P. Muliere
Department of Decision Sciences, Bocconi University, Milan, Italy

BB is able to asymptotically reproduce the underlying dependence structure of that process (Lahiri 2003).

The two most important types of BB are the moving block bootstrap (MBB) of Künsch (1989) and the non-overlapping block bootstrap (NBB) described in Lahiri (2003), based on the blocking rule of Carlstein (1986). The difference between these two schemes is negligible for large sample sizes (Davison and Hinkley 1997; Lahiri 2003). Several generalizations and modifications (e.g. block of blocks) are then possible, and we refer to Lahiri (2003) for more details.

Let $X_1, X_2, \ldots$ be a sequence of stationary random variables and let $\mathcal{X} = \{X_1, \ldots, X_n\}$ indicate the observations. Set $l = l_n \in [1, n]$ to be an integer. Even if it is not strictly necessary for the construction of the MBB, it is frequent to assume that $l \to \infty$ and $n^{-1}l \to 0$ as $n \to \infty$.

Set $\mathcal{B}_i = (X_i, \ldots, X_{i+l-1})$ to be the block of length l that starts with X_i, $1 \leq X_i \leq N = n - l + 1$. The idea is to sample a sufficient number of blocks from $\{\mathcal{B}_1, \ldots, \mathcal{B}_N\}$, in order to obtain the sample $\mathcal{B}_1^*, \ldots, \mathcal{B}_k^*$. If we denote the elements in $\mathcal{B}_i^*$ by $(X_{(i-1)l+1}^*, \ldots, X_{il}^*)$, for $i = 1, \ldots, k$, then $X_1^*, \ldots, X_m^*$ represents the MBB sample[1] of size $m = kl$.

Let us now be interested in the estimator $\hat{\theta}_n = T(F_n)$, where F_n is the empirical distribution function of $X_1, \ldots, X_n$ and $T(\cdot)$ is some real-valued functional. Clearly, $\theta_n^* = T(F_{m,n}^*)$ is the MBB version of $\hat{\theta}_n$, with $F_{m,n}^*$ denoting the empirical distribution of $X_1^*, \ldots, X_m^*$.

The NBB differs from the previous construction only by assuming that the blocks do not overlap. In other words, we consider the largest integer $b \geq 1$ such that $lb \leq n$. Hence $\mathcal{B}_i = (X_{(i-1)l+1}^*, \ldots, X_{il}^*)$, for $i = 1, \ldots, b$.

For all the properties of MBB, NBB and other block bootstrapping techniques we refer to Bing-Yi (1997), Lahiri (2003).

While the NBB was first introduced assuming iid blocks sampled with replacement, for the MBB the assumption of independence does not hold, because of the overlaps. Hence most block bootstrapping techniques assume the exchangeability of blocks, as pointed out in Davison and Hinkley (1997). In fact, if the length of the blocks is sufficiently large,[2] we are able to preserve the dependence structure within blocks and, at the same time, blocks can be assumed exchangeable, i.e. their joint distribution is considered immune to permutations.

Exchangeability (Aldous 1985) is a milestone of Bayesian statistics, since the pioneering works of de Finetti (1931) on his representation theorem. Starting from this consideration, it could be useful to introduce a Bayesian approach to block bootstrapping. This could be done by presenting a particular urn-based sampling mechanism that guarantees the creation of an exchangeable sequence of blocks, on the wake of

---

[1] Because of the moving blocks, in the MBB the first and the last observations of the sample have a lower chance of appearing in the resampling. To avoid this, one can imagine the data points to be placed over a circle with no starting point and no end, according to the so-called end wrapping correction (Davison and Hinkley 1997; Politis and Romano 1992).

[2] Naturally, the choice of blocks' size is a fundamental problem in block bootstrapping techniques; the trade-off between length and number of blocks is well analyzed in Davison and Hinkley (1997). An interesting automatic selection/estimation method for optimal block sizes is given in Politis and White (2004).

what (Lo 1988; Muliere and Secchi 1996; Muliere and Walker 1998; Rubin 1981) have done for single observations. The idea is to use a suitably modified version of the reinforced urn process of Muliere et al. (2000) as sampling scheme.

The paper is organized as follows: in the next section we introduce the urn-based Bayesian block bootstrap; in Sect. 3 we focus our attention on the posterior distribution generated by the UBBB using some simulations; in Sect. 4 we then show a first simple application of the new sampling technique to hydrological data already used to test block bootstrapping techniques; and Sect. 5 actually concludes.

2 The urn-based Bayesian block bootstrap

The sampling mechanism we are about to propose is an adaptation of the reinforced urn process (RUP) of Muliere et al. (2000). RUP constitute a particular class of models that is now widely used in Bayesian nonparametric statistics. A RUP can be seen as a random walk on a state-space of urns and it is basically constructed by the means of Polya urns. We refer to Mahmoud (2009) for a complete treatment of Polya urn models.

Imagine we have a discrete set of states $v_0 < v_1 < v_2 < \ldots$ and associate to each of them a Polya urn $U(v_i)$, $i \geq 0$. Every urn has its own initial composition of white and red balls $C(v_i) = (w(v_i), r(v_i))$, for $i \geq 0$.

A reinforced urn process $\{Y_n\}$ on $\{v_0, v_1, \ldots\}$ is defined iteratively as follows: set $Y_0 = v_0$ and, for $n \geq 1$, if $Y_{n-1} = v_i$ sample urn $U(v_i)$. If the sampled ball is white, then Polya-reinforce urn $U(v_i)$ with $s > 0$ white balls and set $Y_n = v_{i+1}$. Otherwise the ball is red, and add s red balls and set $Y_t = v_0$. In other words, starting from urn $U(v_0)$ we sequentially sample from urns in $v_0, v_1, v_2, \ldots$ until a red ball is chosen, then we start again the sequential sampling.

Repeating the iterative sampling several times, we obtain a sequence like this as realization:

$$
\begin{aligned}
\{Y_n\}_{n \geq 1} &= \{v_0, v_1, v_2, v_0, v_1, v_0, v_1, v_2, v_0, v_1, \ldots, v_{10}, v_{11}, v_0, \ldots.\} \\
&= \{0, 1, 2, 0, 1, 0, 1, 2, 0, 1, \ldots, 10, 11, 0, \ldots.\}
\end{aligned}
\tag{1}
$$

In other words, every time a red ball is sampled, a so called 0-block is created, i.e. a sequence of states between two different 0's.

It can be shown (Muliere et al. 2000) that the process $\{Y_n\}$ is partially exchangeable in the sense of Diaconis and Freedman (1980). In other words $\{Y_n\}$ can be expressed as a mixture of Markov chains. For further details see Muliere et al. (2000).

On the basis of these results, let us now introduce the urn-based Bayesian block bootstrap (UBBB).

Let $X_1, X_2, \ldots$ be a sequence of stationary random variables and let $\mathcal{X} = \{X_1, \ldots, X_n\}$ represent a sample. As in the standard block bootstrap, set $\mathcal{B}_i = (X_i, \ldots, X_{i+l-1})$ to be the block of length l that starts with X_i, $1 \leq X_i \leq N = n - l + 1$. Blocks can overlap or not. This is not relevant for our construction. Imagine we can define a total of k blocks.

Consider now k Polya urns: $U_1, \ldots, U_k$. Every urn, for $i = 1, \ldots, k-1$, initially contains r_i^0 red and w_i^0 white balls. The k-th only contains red balls (i.e. $w_k^0 = 0$). Let $s_i, i = 1, \ldots, k$, be the reinforcement of U_i.[3] Finally, set $Y = (Y_1, Y_2, \ldots)$ to be a sequence of random variables. In order to obtain a sample of blocks according to the UBBB, proceed as follows:

Step 1   Start with the first urn, U_1. Sample it and look at the color of the ball. If it is red, then set $Y_1 = \mathcal{B}_1$, reinforce with s_1 red balls (i.e. $r_1^1 = r_1^0 + s_1$ and $w_1^1 = w_1^0$), and start again with U_1 to generate Y_2. If the ball is white, return the ball to the urn together with s_1 white balls (i.e. $w_1^1 = w_1^0 + s_1$ and $r_1^1 = r_1^0$) and move to Step 2.

Step 2   Sample U_2. If the sampled ball is red set $Y_1 = \mathcal{B}_2$ and Polya-reinforce the urn with s_2 red balls (i.e. $r_2^1 = r_2^0 + s_2$ and $w_2^1 = w_2^0$) and restart from U_1. Otherwise, add back s_2 white balls (i.e. $w_2^1 = w_2^0 + s_2$ and $r_2^1 = r_2^0$) and move to the next Step.

Step... ...

Step k   Sample U_k. With probability 1 the sampled ball is red, hence $Y_1 = \mathcal{B}_k$. Reinforce and start again with Y_2.

Repeating this sampling scheme for several times, we obtain the desired sequence of blocks. It is easy to see the analogy between the UBBB and reinforced urn processes. Every time we restart the sampling, the probability of obtaining a given block depends on the past extractions through the Polya reinforcement mechanism of the urns, which updates their compositions. From a Bayesian point of view, $s_i, i = 1, 2, \ldots, k$, represents the power of knowledge updating, hence the choice of its value has an important impact on the sampling mechanism and on how likely an already sampled block is going to be resampled in the future (see Muliere et al. 2000 for the same aspect with RUP's). Larger values of s correspond to assigning a stronger influence to the sampling mechanism, while smaller values represent the choice of giving more importance to the prior belief, which has been expressed through the initial urns' compositions (see the next section).

Theorem 1 *The sequence $Y = (Y_1, Y_2, \ldots)$ generated by the UBBB is exchangeable.*

Proof The theorem can be proved by observing that the probability of, say, $(\mathcal{B}_1, \mathcal{B}_2, \mathcal{B}_8, \mathcal{B}_k, \ldots, \mathcal{B}_1)$ and $(\mathcal{B}_1, \mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_8, \mathcal{B}_k)$ is just the same, i.e. it is immune to permutations.

It is sufficient to notice that

$$P(Y_1 = \mathcal{B}_l) = \frac{r_l^0}{w_l^0 + r_l^0} \prod_{j=0}^{l-1} \frac{w_j^0}{w_j^0 + r_j^0}, \tag{2}$$

---

[3] We would like to stress that it is not necessary that every urn is always reinforced with the same number s of balls.

$$P(Y_{p+1} > \mathcal{B}_l | Y_1, \ldots, Y_p) = \prod_{j=0}^{l} \frac{w_j^p + s_j f_j}{w_j^p + s_j f_j + r_j^p + s_j d_j^l}, \tag{3}$$

where $f_h = \sum_{i=1}^{p} 1_{\{Y_i = \mathcal{B}_h\}}$ and $d_h^l = \sum_{i=1}^{p} 1_{\{Y_i \neq \mathcal{B}_h | h < l\}}$. $\qquad\square$

Remark 2 If reinforcement $s_i = 0$, for $i = 1, \ldots, k$, then the sequence $Y = (Y_1, Y_2, \ldots)$ is made of independent random variables. Moreover, if all the urns have the same composition and $s_i = s$ for $i = 1, \ldots, k$, we are in the simple case of Bernoulli trials. If the blocks do not overlap, the UBBB is simply a NBB.

If the compositions differ, we are in the field of Poisson trials, i.e. "Bernoulli trials" with different probabilities of success (Pepper 1929).

Acting on the initial compositions and the reinforcements of the different urns, it is hence possible to calibrate the UBBB according to one's needs.

Remark 3 The sampling process of the UBBB can be seen as a generalization of the Polya urn.

Corollary 4 *If the random variables $X_1, X_2, \ldots$ form a Markov chain, then the boot-strap sample $X_1^*, \ldots, X_m^*, m = kl$, is partially exchangeable in the sense of Diaconis and Freedman (1980).*

Proof If $X_1, X_2, \ldots$ is a Markov chain, then within each block of observations the Markovianity is preserved. Then observe that a partially exchangeable process in the sense of Diaconis and Freedman (1980) is nothing but a mixture of Markov chains, where the mixture arises from the exchangeability of the urn-based selection of blocks. $\qquad\square$

While it is not necessary for the original data to be a Markov Chain for the UBBB to work, in the special cases in which this form of dependence is present, the UBBB guarantees that the bootstrap sample we create is at least a mixture of Markov chains. Since partial exchangeability is a basic requirement for several Bayesian techniques, this aspect can be very important for any subsequent analysis.

In Muliere et al. (2000) a result similar to the one of Corollary 4 is obtained by imposing a recurrence condition to the reinforced urn process. In particular, assuming special urn compositions, the authors in Muliere et al. (2000) force the RUP to re-visit a given state in a finite number of steps with a strictly positive probability. Such a condition is, by construction, not necessary in the UBBB. In fact, with probability 1 at least in the last urn, U_k, we sample a red ball, guaranteeing that the process restarts from urn U_1. Recurrence, together with Theorem 1, enables the following important result.

Theorem 5 *The law of the de Finetti measure F of the exchangeable block sequence $Y = (Y_1, Y_2, \ldots, Y_l)$ is that of a discrete beta-Stacy process with parameters $\{w_j/s_j, r_j/s_j : j = 1, \ldots, l\}$.*

Proof The proposition is an application of Theorem 3.26 in Muliere et al. (2000), to which we refer for further details. $\qquad\square$

Hence, in the UBBB we explicit a discrete beta-Stacy process prior on the blocks of observations.

The discrete beta-Stacy process, introduced in Muliere and Walker (1998), represents a random distribution over the set of discrete distributions, and it is a special case of neutral to the right (NTTR) process (Doksum 1974). An interesting property of beta-Stacy processes, that directly derives from being a NTTR random distribution, is conjugacy. In the next section we will focus our attention on this point.

From a Bayesian point of view, the possibility of eliciting a beta-Stacy prior over the blocks is really important, since it allows us to incorporate our eventual a priori knowledge about a special discrete distribution of the blocks. In particular, if F is a discrete beta-Stacy process, it is possible to center it on a given discrete distribution G, so that $E[F(\{j\})] = G(\{j\})$, by considering $c_j > 0 \forall j$, and setting

$$w_j = c_j \times (G(\{j\}) \qquad r_j = c_j \times \left(1 - \sum_{i=0}^{j} (G(\{i\}))\right). \qquad (4)$$

Hence G represents the initial guess, say a multinomial distribution, while the c_j's are the so-called strengths of belief, which represent the weight we assign to distribution G. It is worth noticing that assigning a beta-Stacy process centered on G is much more flexible than directly imposing G on the blocks. In the case our prior guess is not correct, the beta-Stacy process prior is definitely less restrictive than assigning a given distribution with probability 1. While, if G is actually a good guess, the updated beta-Stacy process rapidly concentrates its mass on G through sampling. See Muliere and Walker (1998) for further details.

The possibility of eliciting a very flexible prior and the conjugacy property of beta-Stacy processes make the UBBB an interesting sampling scheme for Bayesians.

In the special case in which G is a Dirichlet distribution $D(w_1^0, r_1^0, \ldots, w_k^0, r_k^0)$, the beta-Stacy process of Theorem 5 reduces to famous Dirichlet process of Ferguson (1973), whose many properties are well known.

For a stationary sequence $X_1, X_2, \ldots, X_N$ and a sample $\mathcal{X} = \{X_1, \ldots, X_n\}$, a Bayesian is interested in the posterior distribution of some estimator $\hat{\theta} = T(H_N, F_n)$ given the data, with H_N indicating the empirical distribution putting $1/N$ on each observation of the whole population ($N \to \infty$), and F_n once again representing the empirical distribution of the sample. In other words, one wants to elicit $\mathcal{L}\{\hat{\theta} = T(H_N, F_n)|F_n\}$.

With the UBBB we aim to approximate the posterior distribution using $\mathcal{L}\{\hat{\theta} = T(F_{m,n}^*, F_n)|F_n\}$, where $F_{p,n}^*$ is a simulated random distribution from a simulation of size $m = N - n$. The procedure is as follows:

1. Generate a sample of size $m = kl$ using the UBBB ($N \approx n + m$);
2. Let $F_{m,n}^* = (nF_n + mG_m^*)/N$, with G_m^* representing the empirical distribution of the sample generated using the UBBB. Evaluate $\hat{\theta} = T(F_{m,n}^*, F_n)$;
3. Repeat the previous step for D times to get $\hat{\theta}_1, \ldots, \hat{\theta}_D$.
4. Use the empirical distribution of $\hat{\theta}_1, \ldots, \hat{\theta}_D$ to approximate the posterior distribution of $\mathcal{L}\{\hat{\theta} = T(H_N, F_n)|F_n\}$.

From this we understand that the UBBB is in some way a generalization of the Polya bootstrap of Lo (1988), in which we sample blocks instead of single observations.

3 The posterior

As we have anticipated, one of the main characteristic of beta-Stacy processes is conjugacy. This means that, in our construction, the posterior distribution for F, after sampling p different blocks, is still a discrete beta-Stacy process, whose parameters have been updated according to the sampling history. In particular we need to count how many times a given block has been sampled so far. For more details we refer to Muliere and Walker (1998).

All in all the bootstrap posterior is a beta-Stacy whose parameters are updated by f_j and d_j^l, for every j, according to Eq. 3. Clearly the behavior of the posterior distribution is influenced by the initial settings of the UBBB (urn compositions, reinforcement, number of urns, etc.), that is to say by the prior knowledge we elicit.

Let us now consider some simulations, in order to analyze the behavior of the posterior distribution using different configurations of parameters and a simple data structure.

We have simulated a basic autoregressive process of order 1, i.e.

$$Z_t = a + \rho Z_{t-1} + \epsilon_t, \tag{5}$$

with $a = 5$, $\rho = 0.8$, $Z_0 = 5$ and $\epsilon_t \sim N(0, 1)$.

We have considered two distinct situations: a time series made of only 30 observations and a time series with 200 time points.

In the first case the estimates we get are $\hat{a} = 2.88(1.00)$ for the intercept and $\hat{\rho} = 0.67(0.12)$ for the autoregressive part. For the second regression we have $\hat{a} = 1.90(0.39)$ and $\hat{\rho} = 0.81(0.04)$. Naturally the number of available observations has a clear impact on our estimates. In brackets we give the standard errors.

For what concerns our prior knowledge of the phenomenon, we analyze two different possibilities:

- in the first case, that we call "No prior", we do not have any prior belief and we consider a sequence of urns with the same initial composition, i.e. containing the same amount of white and red balls, $w_i^0 = r_i^0 = 50 \forall i < k$, where k is the last urn with 1 red ball and 0 white balls, as expected by the UBBB construction. Reinforcement s_i is set equal to $s = 1$ for every urn U_i, $i = 1, \ldots, k$;
- in the second case, called "Prior"[4], we naively give more weight to the first blocks of observations, i.e. we increase the probability of resampling them, by setting $r_i > r_{i+1}$, $i = 1, \ldots, k$, and keeping the total number of balls equal to 100 for all urns (apart from the last one, as usual). Moreover we set $s_i = s = 5$ for every

---

[4] The "Prior" assumption we are introducing must be simply considered as an example and not as the real elicitation of some expert prior knowledge. In particular we are expressing an explicitly wrong belief. A correct prior elicitation can be done using Eq. (4).

Table 1 Parameters' initializations: TS represents the time series; BL is the number of non-overlapping blocks considered; (w_i, r_i) is the composition of urn U_i if available; s is the reinforcement, assumed the same for all urns

TS	BL	(w_1, r_1)	(w_2, r_2)	(w_3, r_3)	(w_4, r_4)	(w_5, r_5)	(w_6, r_6)	(w_7, r_7)	(w_8, r_8)	s
30	2	(50, 50)	(0, 1)	*	*	*	*	*	*	1
30	3	(30, 70)	(40, 60)	(0, 1)	*	*	*	*	*	1
30	2	(50, 50)	(0, 1)	*	*	*	*	*	*	5
30	3	(30, 70)	(40, 60)	(0, 1)	*	*	*	*	*	5
200	4	(50, 50)	(50, 50)	(50, 50)	(0 ,1)	*	*	*	*	1
200	8	(50, 50)	(50, 50)	(50, 50)	(50, 50)	(50, 50)	(50, 50)	(50, 50)	(0, 1)	1
200	4	(20, 80)	(30, 0)	(40, 60)	(0, 1)	*	*	*	*	5
200	8	(20, 80)	(25, 75)	(30, 70)	(35, 65)	(40, 60)	(45 , 55)	(50, 50)	(0, 1)	5

For every parameter specification we have generated 1,000 samples

urn U_i. The increase in the reinforcement can be seen as a countermeasure that balances our prior elicitation.

Remark 6 The choice of s_i is particularly relevant for the finite sample performances of the UBBB.

For example, choosing $s = 1$ implies a quite moderate knowledge update, meaning that the more a block is sampled, the more likely it will be sampled in the future (and its information/dependence structure will be more replicated in the data). The choice $s = 5$ increases the importance of sampling with respect to the prior elicitation given by the initial compositions of the urns.

To reduce the impact of updating one could choose smaller values of s, such as 0.1 or 0.01, moving towards the IID case. On the contrary, setting say $s = 1,000$ causes the first blocks to be sampled over and over again.

Naturally, the impact of s depends on the relative initial composition of the different urns. Adding 1 extra ball to a total of 100 balls is not the same as adding 1 ball to a group of 3 balls. If we standardize the composition of the urns to 1, some reference values for s may be 0.001 (weak reinforcement), 0.01, 0.1 and 0.2 (strong reinforcement).

A nice discussion on the role of reinforcement in reinforced urn processes can be found in Muliere et al. (2000), and it can be easily applied to the UBBB as well.

Table 1 contains all the information concerning the initialization of the UBBB in the different situations. For example we can read that, for the shorter time series made of 30 observations, we have created bootstrap samples using 2 and 3 non-overlapping blocks.[5] Table 1 also tells us that, in the first case, we have initialized the two urns with (50,50) and (0,1) white and red balls.

Under every specification we have created 1,000 bootstrap samples.

---

[5] Please notice that the use of 2 non-overlapping blocks, in a time series containing 30 points, is unrealistic in actual applications. We are proposing it just for pedagogical purposes.
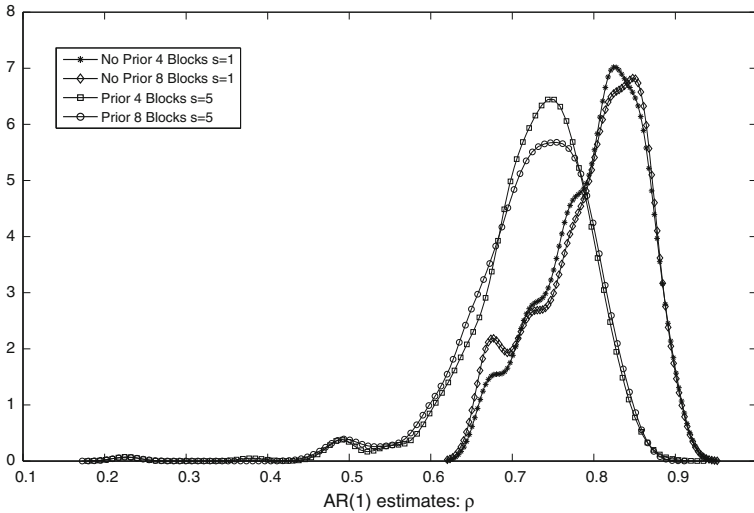
Fig. 1 Posterior distributions for the autoregressive parameter ρ for the longer time series (200 observations) using 1,000 UBBB samples for the 4 specifications of Table 1, rows: 5–8

Figure 1 shows the four posterior distributions for the parameter ρ of the AR(1) model of the longer time series with 200 observations. First of all, we can observe that eliciting some prior knowledge of the phenomenon under investigation has a clear impact on the posterior distribution generated by the UBBB. In particular, while in the case of no specific assumption we have that the posterior distribution tends to concentrate on the true value of $\rho = 0.8$, when we wrongly give more weight to the first blocks (those whose corresponding urns have a higher proportion of red balls) we have that the posterior distribution of ρ tends to slightly shift to the left. This is surely due to the fact that we are giving more strength to the information contained in the first blocks, without actually knowing if this information is correct or not. Eliciting a more appropriate prior knowledge could easily concentrate the posterior on the true value.

Always in Fig. 1 it is interesting to notice that the number and thus the length of non-overlapping blocks has a smaller impact than what one could expect. In particular the length of blocks seems to be important if we elicit a wrong prior. In that case, the larger amount of information about the dependence structure in the data is able to mitigate the effect of a wrong a priori. This is clear if we compare the posterior distribution of ρ under the prior assumption with 4 (50 observation each) and 8 (25 observations) blocks.

Summarizing we get the following basic descriptive statistics about the UBBB posterior distribution of ρ:

- the mean value of ρ is 0.8013 under the no prior assumption with 4 blocks, 0.7991 under the no prior assumption with 8 blocks, 0.7244 under the prior assumption with 4 blocks and 0.7219 under the prior assumption with 8 blocks.

- the corresponding standard deviations are: 0.0510, 0.0621, 0.0764 and 0.5046.

Similar results do hold for the intercept a of the AR(1) model.

Figure 2 contains the posterior distributions for the autoregressive parameter ρ under the four different specifications for the shorter time series with 30 observations. Even in this case it is interesting to notice the impact of the prior specification and the number of blocks. In particular, it is worth noticing how choosing just two blocks made of 15 observations and giving more weight to the first one leads us to a very irregular posterior distribution. In fact we are creating UBBB samples in which the first block (50% of all original data) is more and more present. Having just few observations exacerbates the situation.

For what concerns the mean and the standard deviation of ρ under the UBBB posteriors we get:

- the mean value of ρ is 0.6596 under the no prior assumption with 2 blocks, 0.6323 under the no prior assumption with 3 blocks, 0.6776 under the prior assumption with 2 blocks and 0.6182 under the prior assumption with 3 blocks.

- the corresponding standard deviations are: 0.1155, 0.1304, 0.1084 and 0.1402.

These values are all very close to the estimate we can obtain from the original time series by fitting an AR(1) model, i.e. $\hat{\rho} = 0.67(0.12)$.

Finally, Fig. 3 shows the comparison of the joint distribution of the whole time series with 200 observations, as obtained using the original data set, and those of two bootstrap copies with 4 and 8 blocks, under the "no prior" assumption. Clearly the sample based on larger blocks is definitely closer to the original data. Something similar holds for the 30 observations case, even if the use of a smaller data set presents some more problems.
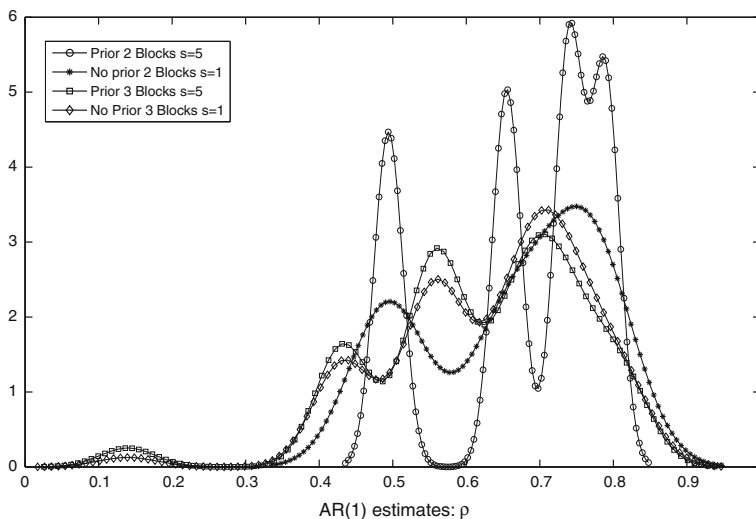
Fig. 2 Posterior distributions for the autoregressive parameter ρ for the shorter time series (30 observations) using 1,000 UBBB samples for the 4 specifications of Table 1, rows: 1–4
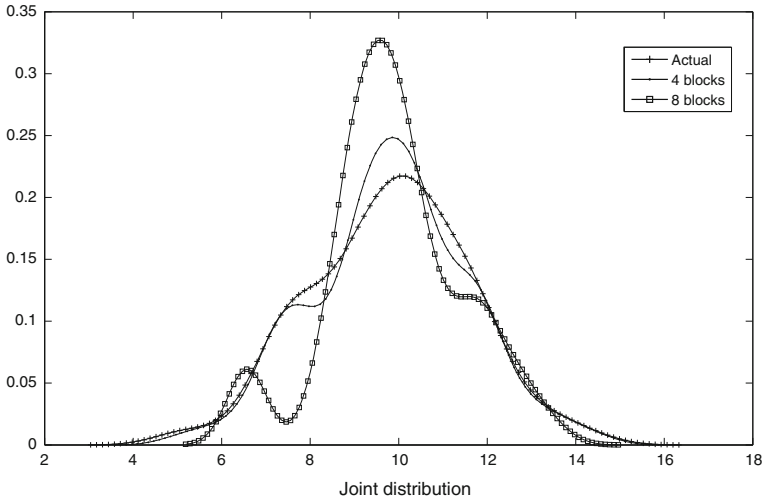
Fig. 3 Comparison of the joint distribution of the longer time series (200 observations) with two UBBB samples with 4 (50 obs.) and 8 blocks (25 obs)

4 An example

We now give a simple demonstration of the UBBB using a time series of river stages from the Rio Negro data set of Brillinger (1988).

The data set contains 120 monthly observations, from January 1970 to December 1979, centered on their mean. This data set has also been used in Davison and Hinkley (1997) to compare different block bootstrapping methods. Figure 4 shows the original series and two examples of UBBB samples with different block lengths.
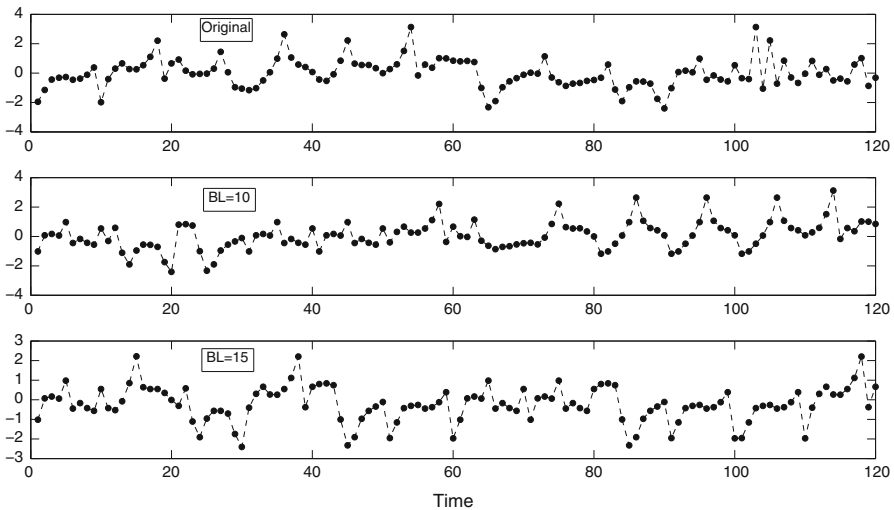
Fig. 4 Resamples from the Rio Negro data. The first plot represent the actual data. Panels 2 and 3 are related to UBBB samples with block length 10 and 15

To evaluate the UBBB we have generated 1,000 samples using different block sizes and no end correction. In particular we have chosen blocks made of 10, 15 and 20 observations (hence $k = 12, 8, 6$). Moreover $s_i = 1, w_i^0 = b_i^0 = 10$ for $i = 1, \ldots, k - 1, w_k^0 = 0$ and $b_k^0 = 1$. These settings guarantee a fair process of reinforcement.

Following (Davison and Hinkley 1997) we have computed the averages (ρ_1, ρ_2 and ρ_3) of the first three autocorrelation coefficients and the SE of the mean.

In the original Rio Negro data, we have $\rho_1 = 0.87(0.003), \rho_2 = 0.57(0.005)$ and $\rho_3 = 0.28(0.012)$, indicating a rather dependent series. For what concerns the mean, the sampling SE is 0.013.

As expected, the larger the blocks used, the more the dependence structure is preserved and the original information replicated. If we start with blocks of 10 observations, we obtain $\hat{\rho}_1 = 0.55, \hat{\rho}_2 = 0.41$ and $\hat{\rho}_3 = 0.18$. The SE of the mean is estimated as 0.28.

Choosing 15 observations per block, the results get better. In fact: $\hat{\rho}_1 = 0.67, \hat{\rho}_2 = 0.50, \hat{\rho}_3 = 0.24$ and SE = 0.32.

Finally, for a block size of 20, we obtain $\hat{\rho}_1 = 0.86, \hat{\rho}_2 = 0.56, \hat{\rho}_3 = 0.24$ and SE = 0.29.

All in all, the UBBB seems to be able to maintain the dependence structure of the original data. Naturally the results improve with the size of the blocks, as expected.

To complete the analysis we have compared our results with those obtainable with other block bootstrap methods, commonly used in the literature (see Davison and Hinkley 1997). In particular, we have tested the UBBB against the stationary bootstrap, the blockwise approach, the blocks of blocks method and a model-based bootstrap, using AR(1) and AR(2) specifications as underlying models. The last two techniques are widely used in the literature thanks to their generally satisfactory results.

Table 2 contains the results: from one side, the UBBB is more accurate than the stationary and the blockwise bootstraps; while, from the other, it seems to be less accurate than the block of blocks and the model-based approaches, in particular when we consider blocks of length 10.

The performances of the UBBB could be improved by better eliciting some prior knowledge about the dependence structure, instead of assuming almost equiprobable blocks. In any case, we believe that these first results[6] are definitely promising and suggest us to continue with the development of the UBBB approach.

5 Conclusions

Block bootstrapping techniques have been introduced in the literature for resampling dependent data, and in particular stationary processes. One of the main assumptions in block bootstrapping is that the blocks of observations are exchangeable.

---

[6] In particular, we would like to underline that the difference between the UBBB and the powerful model-based bootstrap is not extreme, especially if we consider that the AR(2) model is actually the best one for the Rio Negro time series.

Table 2 Comparison of the UBBB with other block bootstrap approaches

Method	Actual →	0.013	0.87	0.57	0.28
	Details	SE	$\hat{\rho}_1$	$\hat{\rho}_2$	$\hat{\rho}_3$
UBBB	10	0.28	0.55	0.41	0.18
	20	0.29	0.86	0.56	0.24
Stationary	10	0.30	0.49	0.40	0.15
	20	0.27	0.79	0.54	0.22
Blockwise	10	0.35	0.53	0.39	0.18
	20	0.28	0.81	0.52	0.22
Blocks of blocks	10	0.31	0.81	0.60	0.39
	20	0.32	0.84	0.53	0.26
Model-based	AR(1)	0.34	0.83	0.59	0.33
	AR(2)	0.45	0.87	0.58	0.30

The first line contains the actual estimates from Rio Negro data. The column "Details" indicates the size of the blocks or the underlying model (model-based bootstrap) used in the experiment

Starting from this assumption, and considering that exchangeability represents a fundamental aspect of Bayesian statistics, we have introduced a new Bayesian approach to block bootstrap using reinforced urn processes (Muliere et al. 2000). We call this new sampling scheme the urn-based Bayesian block bootstrap (UBBB).

The UBBB generates an exchangeable sequence of blocks, eliciting a very flexible beta-Stacy prior on them. This has nice implications, as we discuss in the paper.

By modifying the reinforcement rules and the urns, the UBBB could also be extended and generalized in several ways. The UBBB might even be combined with other techniques, in order to have better replications of the actual data. An idea could be to use the UBBB on blocks of blocks (Lahiri 2003), or with post-blackening techniques (Davison and Hinkley 1997).

Finally, one could apply the UBBB in extreme value theory (EVT). In fact, it is known that one of the most used approach in EVT is the so called block-maxima approach, in which the extremes are computed over blocks of observations. The UBBB could then be used as sampling mechanism for maxima (or minima). Since the UBBB generates a sequence of exchangeable blocks and the maxima are functions of these blocks, the UBBB would generate a sequence of exchangeable maxima with beta-Stacy prior. We will work on this in the future.

In the last part of the paper we have shown an application of our urn-based block bootstrap to some hydrological data, comparing our results with those obtainable with other commonly used block bootstrap techniques.

References

Aldous DJ (1985) Exchangeability and related topics. In: Lecture notes in mathematics, vol 1117. Springer, Berlin

Bing-Yi J (1997) On the relative performance of the block bootstrap for dependent data. Commun Stat Theory Methods 26:1313–1328

Brillinger DR (1988) An elementary trend analysis of the Rio Negro levels at Manaus, 1903–1985. Brazilian J Probab Stat 2:63–79

Carlstein E (1986) The use of subseries methods for estimating the variance of a general statistic from a stationary time series. Ann Stat 14:1171–1179

Davison AC, Hinkley DV (1997) Bootstrap methods and their applications. Cambridge University Press, Cambridge

de Finetti B (1931) Sul significato soggettivo della probabilità. Fundam Math 17:298–329

Diaconis P, Freedman D (1980) de Finetti's theorem for Markov chains. Ann Probab 8:115–130

Doksum KA (1974) Tailfree and neutral random probabilities and their posterior distributions. Ann Probab 2:183–201

Efron B (1979) Bootstrap methods: another look at the jackknife. Ann Stat 7:1–26

Eggenberger F, Polya G (1923) Über die statistik verketteter Vorgänge. Zeitschrift für Angerwandte Math Mech 1:279–289

Ferguson T (1973) Bayesian analysis of some nonparametric problems. Ann Stat 1:209–230

Künsch HR (1989) The jackknife and the bootstrap for general stationary observations. Ann Stat 17:1217–1261

Lahiri SN (2003) Resampling methods for dependent data. Springer, New York

Lo A (1988) A Bayesian bootstrap for finite populations. Ann Stat 16:1684–1685

Mahmoud HM (2009) Polya urn models. CRC Press, Boca Raton

Muliere P, Secchi P (1996) Bayesian nonparametric predictive inference and bootstrap techniques. Ann Inst Stat Math 48:663–673

Muliere P, Secchi P, Walker SG (2000) Urn schemes and reinforced random walks. Stoch Process Appl 88:59–78

Muliere P, Walker S (1998) Extending the family of Bayesian bootstraps and exchangeable urn schemes. J R Stat Assoc B 60:175–182

Pepper ED (1929) On Poisson's series of trials. Math Ann 101:375–380

Phillips PCB (2010, to appear) Bootstrapping I(1) data. J Econ

Politis DN, Romano JP (1992) A circular block resampling procedure for stationary data. In: Lepage R, Billard L (eds) Exploring the limits of bootstrap. Wiley, New York

Politis DN, White H (2004) Automatic block-length selection for the dependent bootstrap. Econ Rev 23:53–70

Rubin D (1981) The Bayesian bootstrap. Ann Stat 9:130–134

Walker S, Muliere P (1997) Beta-stacy processes and a generalization of the Polya urn scheme. Ann Stat 25:1762–1780