

Bootstrapping sample quantiles of discrete data

Carsten Jentsch · Anne Leucht

Received: 13 May 2014 / Revised: 28 October 2014 / Published online: 20 February 2015
© The Institute of Statistical Mathematics, Tokyo 2015

Abstract Sample quantiles are consistent estimators for the true quantile and satisfy central limit theorems (CLTs) if the underlying distribution is continuous. If the distribution is discrete, the situation is much more delicate. In this case, sample quantiles are known to be not even consistent in general for the population quantiles. In a motivating example, we show that Efron’s bootstrap does not consistently mimic the distribution of sample quantiles even in the discrete independent and identically distributed (i.i.d.) data case. To overcome this bootstrap inconsistency, we provide two different and complementing strategies. In the first part of this paper, we prove that m -out-of- n -type bootstraps do consistently mimic the distribution of sample quantiles in the discrete data case. As the corresponding bootstrap confidence intervals tend to be conservative due to the discreteness of the true distribution, we propose randomization techniques to construct bootstrap confidence sets of asymptotically correct size. In the second part, we consider a continuous modification of the cumulative distribution function and make use of mid-quantiles studied in Ma et al. (Ann Inst Stat Math 63:227–243, 2011). Contrary to ordinary quantiles and due to continuity, mid-quantiles lose their discrete nature and can be estimated consistently. Moreover, Ma et al. (Ann Inst Stat Math 63:227–243, 2011) proved (non-)central limit theorems for i.i.d. data, which we generalize to the time series case. However, as the mid-quantile function fails to be differentiable, classical i.i.d. or block bootstrap methods do not lead to completely satisfactory results and m -out-of- n variants are required here as well. The finite sample

C. Jentsch (✉)

Department of Economics, University of Mannheim, L7, 3-5, 68131 Mannheim, Germany
e-mail: cjentsch@mail.uni-mannheim.de

A. Leucht

Institut für Mathematische Stochastik, Technische Universität Braunschweig, Pockelsstraße 14,
38106 Braunschweig, Germany
e-mail: a.leucht@tu-braunschweig.de

performances of both approaches are illustrated in a simulation study by comparing coverage rates of bootstrap confidence intervals.

Keywords Bootstrap inconsistency · Count processes · Mid-distribution function · m -Out-of- n bootstrap · Integer-valued processes

1 Introduction

Since the seminal work of [Efron \(1979\)](#), bootstrapping has been established as a major tool for estimating unknown finite sample distributions of general statistics. Among others, this method has successfully been applied to construct confidence intervals for sample quantiles of continuous distributions; see e.g. [Serfling \(2002, Chapter 2.6\)](#), [Sun and Lahiri \(2006\)](#) and [Sharipov and Wendler \(2013\)](#) and references therein. In this case, the asymptotic behavior of quantile estimators is well understood. Based on the well-known Bahadur representation, a CLT can then be established for sample quantiles in case of an underlying distribution exhibiting a differentiable cumulative distribution function (cdf) and a positive density at the quantile level of interest. This allows for the application of classical results on the bootstrap to mimic the unknown finite sample distribution.

Quantile estimation has many practical applications for discrete-valued data, too. For instance, [Chen and Lazar \(2010\)](#) use it to analyze epileptic seizure count data. Moreover, it plays a central role in survey analysis, e.g. to report the median age at first marriage or the median customer satisfaction, where the latter is typically categorical data. For an overview of results on bootstrapping sample quantiles in the context of survey data, we refer to [Shao and Chen \(1998\)](#). However, the results obtained there rely on certain smoothness assumptions on the underlying distribution, which do generally not hold true for discrete data. In supply chain management, especially for sporadic demand, quantile estimation is required to develop inventory policies that lead to a prescribed α -service level in the sense of [Tempelmeier \(2000, Sect. 2.1\)](#). Confidence intervals can then be used to determine the uncertainty of these estimates.

However, if the underlying distribution is discrete, this task is much more delicate than in the continuous case. Sample quantiles may not even be consistent in general for the population quantiles in this case. This issue occurs due to the fact that the cdf is a step function. This leads to inconsistency if the level of the quantile of interest lies in the image of the cdf and, consequently, CLTs do not hold true anymore. Before we illustrate this inconsistency with the help of a simple, but very insightful toy example below, first, we fix some notation that is used throughout this paper. Let Q_p for $p \in (0, 1)$ be the usual population p -quantile of a probability distribution with cdf F defined via its generalized inverse, i.e.

$$Q_p = F^{-1}(p) = \inf_t \{t : F(t) \geq p\}. \quad (1)$$

With observations X_1, \dots, X_n at hand, the sample p -quantile \widehat{Q}_p is defined as the empirical counterpart to (1), that is,

$$\widehat{Q}_p = \widehat{F}_n^{-1}(p) = \inf_t \{t : \widehat{F}_n(t) \geq p\}, \tag{2}$$

where $\widehat{F}_n(x) = n^{-1} \sum_{i=1}^n 1(X_i \leq x)$ denotes the empirical distribution. Here and in the sequel, $\lceil x \rceil$ ($\lfloor x \rfloor$) denotes the smallest (largest) integer that is larger (smaller) or equal to x .

Toy example: coin flip data Suppose a coin is flipped independently n times and we observe a sequence X_1, \dots, X_n of zeros and ones such that $P(X_i = 0) = \theta = 1 - P(X_i = 1)$ for some $\theta \in (0, 1)$. Let $X_{med} = Q_{0.5}$ and $\widehat{X}_{med} = \widehat{Q}_{0.5}$ denote the population median and the sample median, respectively. This leads to

$$P(\widehat{X}_{med} = 0) = \sum_{k=\lceil \frac{n}{2} \rceil}^n \binom{n}{k} \theta^k (1 - \theta)^{n-k}. \tag{3}$$

If a coin is fair, i.e. $\theta = 1/2$, we have $X_{med} = 0$ and, by symmetry properties, we get

$$P(\widehat{X}_{med} = 0) = \begin{cases} \frac{1}{2}, & n \text{ odd} \\ \frac{1}{2} + \binom{n}{n/2} \left(\frac{1}{2}\right)^{n+1}, & n \text{ even} \end{cases}. \tag{4}$$

From Stirling’s formula (see e.g. Krantz 1991, Theorem 10.23), we get $\binom{n}{n/2} \left(\frac{1}{2}\right)^{n+1} = O(n^{-1/2})$, which leads to

$$P(\widehat{X}_{med} = 0) = 1 - P(\widehat{X}_{med} = 1) \rightarrow \frac{1}{2} \tag{5}$$

as $n \rightarrow \infty$. This is contrary to $\widehat{X}_{med} \xrightarrow{P} 0$, i.e. the sample median is not a consistent estimator and its limiting distribution is an equally-weighted 2-point distribution.

In this paper, as a first result, we show that one consequence of the estimation inconsistency illustrated in (5) is that the classical bootstrap of Efron for i.i.d. data is inconsistent for sample quantiles if they do not consistently estimate the true quantile. More precisely, we prove that the Kolmogorov-Smirnov distance between the cdf’s and their bootstrap analogues does not converge to zero, but to non-degenerate random variables. These turn out to be functions of a random variable $U \sim \text{Unif}(0, 1)$ in the special case of the sample median for the fair coin flip discussed in the example and in Theorem 1 in Sect. 1. To the authors’ knowledge, such a specific phenomenon has not been observed in the bootstrap literature so far.

Toy example: bootstrapping coin flip data Let X_1^*, \dots, X_n^* be i.i.d. (Efron) bootstrap replicates of X_1, \dots, X_n and let \widehat{X}_{med}^* denote the bootstrap sample median based on the bootstrap observations. Then, we have analogously to (3)

$$P^*(\widehat{X}_{med}^* = 0) = \sum_{k=\lceil \frac{n}{2} \rceil}^n \binom{n}{k} \widehat{\theta}_n^k (1 - \widehat{\theta}_n)^{n-k}, \tag{6}$$

where $\widehat{\theta}_n = n^{-1} \sum_{t=1}^n 1(X_t = 0)$ and P^* denotes as usual the bootstrap distribution (conditional on X_1, \dots, X_n). In Theorem 1 below, we show that

$$P^*(\widehat{X}_{\text{med}}^* = 0) = 1 - P^*(\widehat{X}_{\text{med}}^* = 1) \xrightarrow{\mathcal{D}} U \sim \text{Unif}(0, 1). \tag{7}$$

By combining the result in (7) with (5), we get inconsistency of Efron’s bootstrap, see Theorem 1 below for details.

In view of the results displayed in the toy example, it is worth noting that, more generally, the population p -quantile Q_p may be defined as any real number q that satisfies the two inequalities

$$P(X \leq q) \geq p \quad \text{and} \quad P(X \geq q) \geq 1 - p, \tag{8}$$

where $X \sim F$, i.e. the definition (1) corresponds to the smallest possible value of q in (8). In particular, it is not unusual to define the median X_{med} as the center of the smallest and the largest possible values of the median with respect to definition (8). The sample median \widehat{X}_{med} is then defined in direct analogy. However, this choice does not affect at all the inconsistency results above and we prefer the definitions via (1) and (2) for two reasons. Firstly, they naturally fit into the more general notation of the (generalized) inverse of the cdf and, secondly, the (sample) median then takes values in the support of P^X only.

Still, one would like to establish consistent bootstrap results not only for the continuous setting, but also in general for discrete distributions. In this paper, as the use of ordinary quantiles in discrete settings can be discussed conversely, we provide two different and complementing strategies to tackle the issue of bootstrap inconsistency for sample quantiles in the discrete setup that is illustrated in the toy example above.

In the first part of this paper, we investigate whether the m -out-of- n bootstrap (or low-intensity bootstrap) leads to asymptotically consistent bootstrap approximations. In several contexts where the classical bootstrap fails, this modified bootstrap scheme is known to be a valid alternative; see e.g. [Swanepoel \(1986\)](#), [Angus \(1993\)](#), [Deheuvels et al. \(1993\)](#), [Athreya and Fukuchi \(1994, 1997\)](#) and [Del Barrio et al. \(2013\)](#) among others. We prove that the i.i.d. m -out-of- n bootstrap is consistent for sample quantiles *without centering* in the i.i.d. discrete data case, but also that inconsistency for Efron’s bootstrap remains if the procedure is applied *with centering*. These differing results seem to be odd at first sight, but they can be explained by systematically different centering schemes. Another somewhat surprising result is that, on the one hand, bootstrap consistency can be achieved for i.i.d. data as well as dependent time series data for one and the same i.i.d. m -out-of- n bootstrap procedure (without centering) as long as only *single* sample quantiles are considered. But on the other hand, an m -out-of- n block bootstrap procedure à la [Athreya et al. \(1999\)](#) has to be used to mimic correctly the joint limiting distribution of *several* sample quantiles in the time series case. To be able to establish this theory, we had to derive the joint limiting distribution of vectors of sample quantiles for weakly dependent time series processes. This might be of independent interest.

The consistency results achieved for the m -out-of- n bootstrap are then applied to construct bootstrap confidence intervals. As these tend to be conservative due to the

discreteness of the true distribution, we propose randomization techniques similar to the construction of randomized tests (e.g. classical Neyman–Pearson tests) to construct bootstrap confidence intervals of asymptotically correct size.

All afore-mentioned difficulties related to discrete distributions are mainly due to the jumps occurring in the distribution function, which leads to many quantiles having the same values. Another look at quantiles of discretely distributed data is to employ the so-called mid-distribution function proposed by Parzen (1997, 2004). This concept has been further studied in Ma et al. (2011) and has been applied successfully e.g. to probabilistic index models in Thas et al. (2006). The corresponding mid-quantile function is a continuous, piecewise linear modification of the ordinary quantile function.

In the second part of this paper, we make use of mid-quantiles. Although the distributions of the mid-quantiles lose their discrete nature, they allow for a meaningful interpretation in many relevant situations. Exemplary, compare two (small) samples stemming from coin flip scenarios. Both their sample medians may be computed to 0. Actually, this is not much information since the samples widely may differ. Assume for example that in the first sample five out of nine heads (equal to 0) may be occurred and in the second sample eight out of nine heads occurred. It would be of much more use to regard the empirical proportion of heads and tails within each sample to describe their underlying distributions and to reflect possible differences. Based on such considerations Parzen (1997, 2004) established the concept of mid-distribution functions to handle sample medians more likely. Contrary to ordinary quantiles, it turns out that the mid-quantiles can be estimated consistently. Moreover, (non-)central limit theorems of the sample mid-quantiles can be achieved, where the limiting distributions crucially depend on whether the mid-distribution function is differentiable or not at the quantile of interest.

First, we generalize the limiting results obtained in Ma et al. (2011) to the time series case under a so-called τ -weak dependence condition introduced by Dedecker and Prieur (2005). This extension is motivated by a growing literature on modeling of and statistical inference for count data that appear, e.g. as transaction numbers of financial assets or biology where the evolution of infection numbers over time is of great interest; see for instance Fokianos et al. (2009) and Ferland et al. (2006). In particular, the theory provided in this paper covers parameter-driven integer-valued autoregressive (INAR) models but also observation-driven integer-valued GARCH (INGARCH) models. By construction, the mid-quantile function is continuous, but it fails to be differentiable. Caused by this non-smoothness, it turns out that classical i.i.d. or block bootstrap methods do not lead to completely satisfactory results and m -out-of- n variants are required here as well. Moreover, due to boundary effects, randomization techniques still have to be used to construct confidence intervals of asymptotic correct level $1 - \alpha$ for $\alpha \in (0, 1)$.

The rest of the paper is organized as follows. Section 2 focuses on bootstrapping classical quantiles. In a first Sect. 2.1 we show inconsistency of Efron's bootstrap in the special case of the fair coin flip. Afterwards, in Sect. 2.2 we discuss validity of low-intensity bootstrap methods for quantiles in a much more general framework that covers a large class of discretely distributed time series. In Sect. 2.3 randomization techniques for the construction of confidence sets are provided before the finite

sample behavior of our methods is illustrated in Sect. 2.4. In Sect. 3 we consider the alternative concept of mid-quantiles. In Sect. 3.1 we generalize the asymptotic results established in Ma et al. (2011) for the i.i.d. case to the case of weakly dependent time series data. Bootstrap validity is discussed in Sect. 3.2 and, based on these results, confidence intervals for mid-quantiles are provided in Sect. 3.3. Numerical experiments are reported in Sect. 3.4. Finally, both concepts are discussed in a comparative conclusion. All proofs and auxiliary results are deferred to a final section of the paper.

2 Bootstrapping sample quantiles

2.1 Inconsistency of Efron’s bootstrap

In this section, we prove for the simple example of a fair coin flip and the sample median that Efron’s bootstrap is not capable in general to estimate consistently the limiting distribution of sample quantiles from discretely distributed data. To check for bootstrap consistency, we make use of the Kolmogorov–Smirnov distance and show that neither

$$d_{KS}(\widehat{X}_{med}^*, \widehat{X}_{med}) = \sup_{x \in \mathbb{R}} |P^*(\widehat{X}_{med}^* \leq x) - P(\widehat{X}_{med} \leq x)| \tag{9}$$

(without centering) nor

$$\begin{aligned} d_{KS}(\widehat{X}_{med}^* - \widehat{X}_{med}, \widehat{X}_{med} - X_{med}) \\ = \sup_{x \in \mathbb{R}} |P^*(\widehat{X}_{med}^* - \widehat{X}_{med} \leq x) - P(\widehat{X}_{med} - X_{med} \leq x)| \end{aligned} \tag{10}$$

(with centering) converges to zero for increasing sample size, but to non-degenerate distributions, which turn out to be different in these two cases. Dealing with the non-centered case (9) first and due to $X_i \in \{0, 1\}$ for the coin flip example, it suffices to consider

$$\sup_{x \in \{0,1\}} |P^*(\widehat{X}_{med}^* \leq x) - P(\widehat{X}_{med} \leq x)| = |P^*(\widehat{X}_{med}^* = 0) - P(\widehat{X}_{med} = 0)|, \tag{11}$$

because $|P^*(\widehat{X}_{med}^* \leq x) - P(\widehat{X}_{med} \leq x)| = 0$ holds for all $x \notin \{0, 1\}$. Further, we know that $P(\widehat{X}_{med} = 0) \rightarrow 1/2$ with $n \rightarrow \infty$ by (5) such that we have to investigate

$$P^*(\widehat{X}_{med}^* = 0) = \sum_{k=\lceil \frac{n}{2} \rceil}^n \binom{n}{k} \widehat{\theta}_n^k (1 - \widehat{\theta}_n)^{n-k} \tag{12}$$

in more detail. For the case with centering (10), things become slightly different and it suffices to consider

$$\sup_{k \in \{-1,0\}} |P^*(\widehat{X}_{med}^* - \widehat{X}_{med} \leq k) - P(\widehat{X}_{med} - X_{med} \leq k)| \tag{13}$$

in this case. Precisely, we get the following results.

Theorem 1 (Inconsistency of Efron’s bootstrap) *For independent and fair ($\theta = 0.5$) coin flip random variables X_1, \dots, X_n and i.i.d. (Efron) bootstrap replicates X_1^*, \dots, X_n^* , it holds*

$$P^*(\widehat{X}_{\text{med}}^* = 0) = \sum_{k=\lceil \frac{n}{2} \rceil}^n \binom{n}{k} \widehat{\theta}_n^k (1 - \widehat{\theta}_n)^{n-k} \xrightarrow{\mathcal{D}} U \sim \text{Unif}(0, 1). \tag{14}$$

This leads to:

(i) *For Efron’s bootstrap without centering, it holds*

$$d_{KS}(\widehat{X}_{\text{med}}^*, \widehat{X}_{\text{med}}) \xrightarrow{\mathcal{D}} \left| U - \frac{1}{2} \right| \sim \text{Unif}(0, 1/2). \tag{15}$$

(ii) *For Efron’s bootstrap with centering, it holds*

$$d_{KS}(\widehat{X}_{\text{med}}^* - \widehat{X}_{\text{med}}, \widehat{X}_{\text{med}} - X_{\text{med}}) \xrightarrow{\mathcal{D}} 1 \left(\frac{1}{2} \leq U \right) U + 1 \left(\frac{1}{2} > U \right) - \frac{1}{2} =: S, \tag{16}$$

where the cdf of S is given by

$$F_S(x) = x 1_{[0, \frac{1}{2}]}(x) + 1_{[\frac{1}{2}, \infty)}(x).$$

2.2 The m -out-of- n bootstrap

2.2.1 Coin flip data

Of course, there are other situations discussed in the literature, where the ordinary Efron’s bootstrap fails; see Bickel and Friedman (1981, Section 6), Mammen (1992) and Horowitz (2001) and references therein. The most prominent example is the maximum of i.i.d. random variables X_1, \dots, X_n , that is, $M_n = \max(X_1, \dots, X_n)$. In this case, bootstrap inconsistency of $M_n^* = \max(X_1^*, \dots, X_n^*)$ has been investigated in Angus (1993). To circumvent this problem and in view of the well-known limiting result [cf. Resnick (1987), Chapter 1]

$$P(a_n^{-1}(M_n - b_n) \leq x) \xrightarrow[n \rightarrow \infty]{} G(x) \quad \forall x \in \mathbb{R}$$

for suitable distributions P^{X_1} , sequences $(a_n)_n$ and $(b_n)_n$ and a non-degenerate cdf G , Swanepoel (1986), Deheuvels et al. (1993) and Athreya and Fukuchi (1994, 1997) used the low-intensity m -out-of- n -bootstrap. That is, drawing with replacement m times with $m \rightarrow \infty$ such that $m = o(n)$ to get X_1^*, \dots, X_m^* and to mimic the distribution of $a_n^{-1}(M_n - b_n)$ by that of $a_m^{-1}(M_m^* - b_m)$. This task has been generalized by Athreya

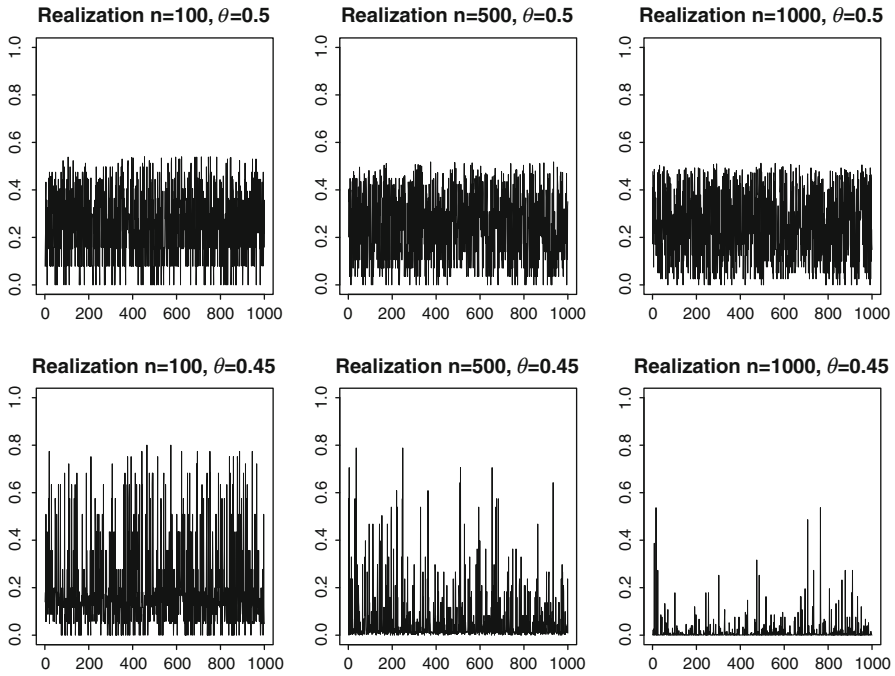


Fig. 1 1000 realizations of $\sup_k |P^*(\widehat{X}_{\text{med}}^* \leq k) - P(\widehat{X}_{\text{med}} \leq k)|$ of a coin flip for sample sizes $n \in \{100, 500, 1000\}$ (from left to right) and for $\theta \in \{0.5, 0.45\}$ (from top to bottom)

et al. (1999) to time series data, where additionally a low-intensity block bootstrap has been proposed and investigated.

The situation addressed in this paper is somehow comparable. A closer inspection of (3) and (6) leads to the conclusion that if we were allowed to replace $\widehat{\theta}_n$ by θ for asymptotic considerations, we would get the same limiting results. Obviously, from (5) and (7), this is not the case. However, as

$$\sqrt{n}(\widehat{\theta}_n - \theta) \xrightarrow{D} \mathcal{N}(0, \theta(1 - \theta)), \tag{17}$$

inconsistency stated in Theorem 1 for the coin flip can be explained by the fact that the convergence $\widehat{\theta}_n - \theta = O_P(n^{-1/2})$ is just “too slow”. Hence, the bootstrap is not able to mimic the underlying scenario correctly since the latter completely differs for $\theta = 1/2$ and $\theta \neq 1/2$. Note that the limiting distribution is a non-degenerate 2-point distribution in the first and degenerate in the second case; compare Theorem 6 below and Fig. 1. Therefore, natural questions are whether an m -out-of- n bootstrap may be capable to “speed up” the convergence of $\widehat{\theta}_n$ (relative to the convergence of the empirical cdf on the bootstrap side) and whether this does lead to bootstrap consistency. The following theorem summarizes our findings in this direction for the sample median without and with centering corresponding to the results of Theorem 1.

Theorem 2 (Consistency and inconsistency for the m -out-of- n bootstrap for the sample median) *For independent and fair ($\theta = 0.5$) coin flip random variables X_1, \dots, X_n , we draw i.i.d. bootstrap replicates X_1^*, \dots, X_m^* . Suppose that $m/n + 1/m = o(1)$ as $n \rightarrow \infty$ and denote the bootstrap sample median based on X_1^*, \dots, X_m^* by $\widehat{X}_{m,med}^*$.*

(i) *For the m -out-of- n bootstrap without centering, it holds*

$$d_{KS}(\widehat{X}_{m,med}^*, \widehat{X}_{med}) \xrightarrow{P} 0.$$

(ii) *For the m -out-of- n bootstrap with centering, it holds*

$$d_{KS}(\widehat{X}_{m,med}^* - \widehat{X}_{med}, \widehat{X}_{med} - X_{med}) \xrightarrow{D} \frac{1}{2} \mathbf{1} \left(U < \frac{1}{2} \right) =: \widetilde{S},$$

where $U \sim \text{Unif}(0, 1)$ such that $2\widetilde{S} \sim \text{Bin}(1, 0.5)$ is Bernoulli-distributed.

Remark 3 The results of Theorem 2 that state consistency for the non-centered sample median, but inconsistency for the centered version for the m -out-of- n bootstrap, seem to be surprising at first sight. However, by a closer inspection of part (ii) this oddity can be explained by the fact that $X_{med} = 0$, while \widehat{X}_{med} and $\widehat{X}_{m,med}^*$ take the values 0 and 1 with limiting probability 1/2 each. Hence, the centering differs on the bootstrap and the non-bootstrap side of (ii). This effect is caused by the estimation inconsistency of the sample median.

In Fig. 2, the differing asymptotic behavior of $\widehat{X}_{med,m}^*$ for $m = n$ and $m = n^{2/3}$ is illustrated via histogram plots for coin flip data. For the first case, the asymptotic uniform distribution of $P^*(\widehat{X}_{med,n}^* = 0)$ is reflected by the high variability of the histograms, whereas the probabilities seem to be more balanced in the second case.

So far we have considered only the toy example of i.i.d. fair coin flip random variables and the sample median. This seems to be very restrictive at first sight. In the following, we turn to a much more general setup and show that asymptotics follow immediately from the results established for the coin flip example. Consequently, it turns out to be not that toyish at all.

2.2.2 General setup

We now turn to more general distributions than the Bernoulli distribution and suppose that $(X_t)_{t \in \mathbb{Z}}$ is a sequence of random variables that might inherit a certain dependence structure. In the last decade, Poisson autoregressions [e.g. Ferland et al. (2006) and Fokianos et al. (2009)], INAR processes [e.g. McKenzie (1988), Weiß (2008) and Drost et al. (2009)] and various extensions of these models have attracted increasing interest, see Fokianos (2011). We intend to derive results that hold true for a broad range of processes including the previous one. Doukhan et al. (2012a, b) showed that these processes are τ -dependent with geometrically decaying coefficients. Therefore, we will use this concept in the sequel and state its definition for sake of completeness.

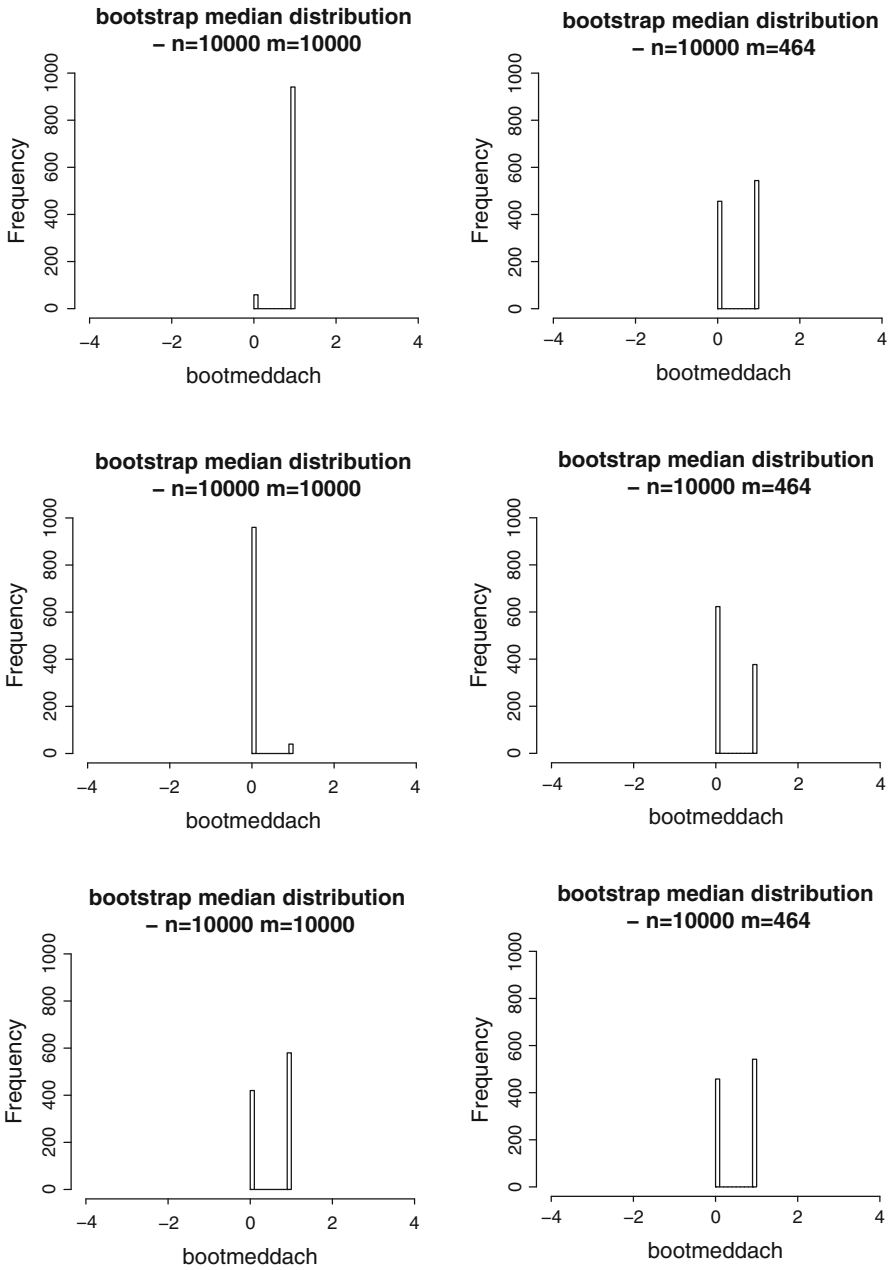


Fig. 2 Histograms of $\hat{X}_{\text{med},m}^*$ based on i.i.d. bootstrap replicates X_1^*, \dots, X_m^* from fair coin flip data X_1, \dots, X_n for $n = 10000$ and $m = n$ (first column) and $m = n^{2/3}$ (second column)

However, it can be seen from the proofs below that any other concept of weak dependence being sufficient for a CLT of the empirical distribution function can be applied here as well.

Definition 4 Let (Ω, \mathcal{A}, P) be a probability space and $(X_t)_{t \in \mathbb{Z}}$ be a strictly stationary sequence of integrable \mathbb{R}^d -valued random variables. The process is called τ -(weakly) dependent if

$$\tau(h) = \sup_{D \in \mathbb{N}} \frac{1}{D} \sup_{h \leq t_1 < \dots < t_D} \left\{ \tau(\sigma(X_t, t \leq 0), (X_{t_1}, \dots, X_{t_D})) \right\} \xrightarrow{h \rightarrow \infty} 0,$$

where

$$\tau(\mathcal{M}, X) = E \left(\sup_{f \in \Lambda_1(\mathbb{R}^p)} \left| \int_{\mathbb{R}^p} f(x) dP^{X|\mathcal{M}}(x) - \int_{\mathbb{R}^p} f(x) dP^X(x) \right| \right).$$

Here, \mathcal{M} is a sub- σ -algebra of \mathcal{A} , $P^{X|\mathcal{M}}$ denotes the conditional distribution of the \mathbb{R}^p -valued random variable X given \mathcal{M} , and $\Lambda_1(\mathbb{R}^p)$ denotes the set of 1-Lipschitz functions from \mathbb{R}^p to \mathbb{R} , i.e. $f \in \Lambda_1(\mathbb{R}^p)$ if $|f(x) - f(y)| \leq \|x - y\|_1 = \sum_{j=1}^p |x_j - y_j| \forall x, y \in \mathbb{R}^p$.

Remark 5 If a process $(X_t)_{t \in \mathbb{Z}}$ on (Ω, \mathcal{A}, P) is τ -dependent and if \mathcal{A} is rich enough, then there exists, for all $t < t_1 < \dots < t_D \in \mathbb{Z}$, $D \in \mathbb{N}$, a random vector $(\tilde{X}_{t_1}, \dots, \tilde{X}_{t_D})'$ which is independent of $(X_s)_{s \leq t}$, has the same distribution as $(X_{t_1}, \dots, X_{t_D})'$ and satisfies

$$\frac{1}{D} \sum_{j=1}^D E \|\tilde{X}_{t_j} - X_{t_j}\|_1 \leq \tau(t_1 - t); \tag{18}$$

cf. [Dedecker and Prieur \(2004\)](#). This L_1 -coupling property will be an essential device for the proofs of our results below. Also note that in particular sequences of i.i.d. random variables $(X_t)_{t \in \mathbb{Z}}$ are τ -dependent with $\tau(0) \leq 2E\|X_1\|$ and $\tau(h) = 0$ for $h \neq 0$. Nevertheless, we state the i.i.d. case separately in all our Theorems since τ -dependent processes are assumed to have finite first moment which is not necessary in our results if the data are i.i.d..

Regarding the marginal distribution P^{X_1} , we assume that it has support $\text{supp}(P^{X_1}) = V$, that is, $P(X_i \in V) = 1$, where

$$V = \{v_j \mid j \in T \subseteq \mathbb{Z}\} \tag{19}$$

for some finite or countable index set T with $v_j < v_{j+1}$ for all $j \in T$. Further, we assume that V has no accumulation point. As the cdf F is a step function, there is always a $p \in (0, 1)$ such that the p -quantile $Q_p = v_j$, say, as well as v_{j+1} satisfies both inequalities in (8). Recall that this covers particularly the population median in

the fair coin flip example. In the following, we consider the asymptotics for the sample quantile \widehat{Q}_p as defined in (2) and its bootstrap analogue

$$\widehat{Q}_{p,m}^* = (\widehat{F}_m^*(p))^{-1} = \inf_t \{t : \widehat{F}_m^*(t) \geq p\},$$

where $\widehat{F}_m^*(x) = m^{-1} \sum_{i=1}^m 1(X_i^* \leq x)$ denotes the empirical bootstrap distribution function. Similar to (3), for all $x \in \mathbb{R}$, we have

$$P(\widehat{Q}_p \leq x) = P\left(\sum_{i=1}^n 1(X_i \leq x) \geq \lceil np \rceil\right) = \sum_{j=\lceil np \rceil}^n \binom{n}{j} F^j(x)(1 - F(x))^{n-j}.$$

For the bootstrap p -quantile $\widehat{Q}_{p,m}^*$ based on i.i.d. bootstrap pseudo replicates X_1^*, \dots, X_m^* , we get the analogue representation

$$\begin{aligned} P^*(\widehat{Q}_{p,m}^* \leq x) &= P^*\left(\sum_{i=1}^m 1(X_i^* \leq x) \geq \lceil mp \rceil\right) \\ &= \sum_{j=\lceil mp \rceil}^m \binom{m}{j} \widehat{F}_n^j(x)(1 - \widehat{F}_n(x))^{m-j}. \end{aligned}$$

Further, for all $x \in \mathbb{R}$ and analogue to (17), we have for the i.i.d. case

$$\sqrt{n}(\widehat{F}_n(x) - F(x)) \xrightarrow{\mathcal{D}} \mathcal{N}(0, W) \quad \text{where } W = \text{cov}(1(X_0 \leq x), 1(X_0 \leq x)).$$

As for the median in the coin flip example and analogue to (9) and (10), to check for bootstrap consistency, we have to consider $d_{KS}(\widehat{Q}_p^*, \widehat{Q}_p)$ or $d_{KS}(\widehat{Q}_p^* - \widehat{Q}_p, \widehat{Q}_p - Q_p)$. To this end, we first study the asymptotics for the empirical quantile \widehat{Q}_p . In particular, part (iii) of the following lemma addresses the joint limiting distributions of several empirical quantiles. To the authors knowledge, such a result has not been established in this generality so far and may be of independent interest.

Theorem 6 (Asymptotics of empirical quantiles for discrete distributions) *Let X_1, \dots, X_n be discretely distributed random variables which are either i.i.d. or observations of a strictly stationary and τ -dependent process $(X_t)_{t \in \mathbb{Z}}$ with $\sum_{h=0}^\infty \tau(h) < \infty$ and $\text{supp}(P^{X_1}) = V$ as described above.*

(i) *If $F(Q_p) > p$,*

$$P(\widehat{Q}_p = Q_p) \xrightarrow[n \rightarrow \infty]{} 1.$$

(ii) *If $F(Q_p) = p$ and $Q_p = v_j$, say, for some $v_j \in V$,*

$$P(\widehat{Q}_p = v_j) \xrightarrow[n \rightarrow \infty]{} 1/2 \quad \text{and} \quad P(\widehat{Q}_p = v_{j+1}) \xrightarrow[n \rightarrow \infty]{} 1/2.$$

(iii) For p_1, \dots, p_d such that $F(Q_{p_i}) = p_i, i = 1, \dots, k$ and $F(Q_{p_i}) > p_i, i = k + 1, \dots, d$ with $Q_{p_i} = v_{i_i}$, say, joint convergence in distribution of $\widehat{Q} = (\widehat{Q}_{p_1}, \dots, \widehat{Q}_{p_d})'$ holds. Precisely, we have

$$P(\widehat{Q} = \underline{q}) \xrightarrow{n \rightarrow \infty} \begin{cases} P\left(\bigcap_{j=1}^k \{2 \cdot 1(q_j = Q_{p_j}) - 1\} Z_j \geq 0\right), & q_i = Q_{p_i}, i = k + 1, \dots, d \\ 0, & \text{otherwise} \end{cases} \quad (20)$$

where $\underline{q} = (q_1, \dots, q_d)'$ with $q_i \in \{v_{i_i}, v_{i_i+1}\}$. Here, the probability of the empty intersection is set to one and $\underline{Z} = (Z_1, \dots, Z_k)' \sim \mathcal{N}(\underline{0}, \mathbf{W})$ with covariance matrix \mathbf{W} having entries

$$W_{i,j} = \begin{cases} \text{cov}(1(X_0 \leq q_i), 1(X_0 \leq q_j)), & \text{i.i.d. case} \\ \sum_{h \in \mathbb{Z}} \text{cov}(1(X_h \leq q_i), 1(X_h \leq q_j)), & \text{time series case} \end{cases}$$

Note that the asymptotics do not depend on the dependence structure of the underlying process as long as single quantiles are considered, compare Remark 10. This does no longer hold true when the joint distribution of several quantiles is considered. Part (iii) above shows that \widehat{Q} converges to a random variable with 2-point marginal distributions that are indeed *dependent* not only for the time series case, but also for i.i.d. random variables. More precisely, the probability that the vector of empirical quantiles \widehat{Q} equals the vector \underline{q} corresponds asymptotically to the probability that the normally distributed random variable \underline{Z} takes values in a certain orthant of \mathbb{R}^k depending on \underline{q} . This is illustrated in the following example.

Example 7 In the situation of Theorem 6(iii) let $k = 2$ and suppose $(Q_{p_1}, Q_{p_2}) = (v_{i_1}, v_{i_2})$.

- (i) If $\underline{q} = (v_{i_1}, v_{i_2})$, we have $P(\widehat{Q} = \underline{q}) \xrightarrow{n \rightarrow \infty} P(0 \leq Z_1, 0 \leq Z_2)$.
- (ii) If $\underline{q} = (v_{i_1}, v_{i_2+1})$, we have $P(\widehat{Q} = \underline{q}) \xrightarrow{n \rightarrow \infty} P(0 \leq Z_1, 0 \geq Z_2)$.

After having established asymptotic theory for sample quantiles in this general setup, it remains to consider the bootstrap analogue, i.e. $P^*(\widehat{Q}_{p,m}^* \leq x)$, in more detail. In particular, for $x = Q_p$ we have by Theorem 8 below

$$\begin{aligned} P^*(\widehat{Q}_{p,m}^* = Q_p) &= P^*(\widehat{Q}_{p,m}^* \leq Q_p) - o_P(1) \\ &= \sum_{k=\lceil mp \rceil}^m \binom{m}{k} \widehat{F}_n^k(Q_p) (1 - \widehat{F}_n(Q_p))^{m-k} - o_P(1), \end{aligned}$$

which has (asymptotically) exactly the same shape as (12), and the results of Theorems 1 and 2 transfer directly to this more general setup.

Theorem 8 (Consistency of the i.i.d. m -out-of- n bootstrap) *Let X_1, \dots, X_n be discretely distributed i.i.d. random variables with $\text{supp}(P^{X_1}) = V$ as above and we draw i.i.d. bootstrap replicates X_1^*, \dots, X_m^* . Suppose that $m/n + 1/m = o(1)$ as $n \rightarrow \infty$*

and let $\widehat{Q} = (\widehat{Q}_{p_1}, \dots, \widehat{Q}_{p_d})$ as in Theorem 6 and $\widehat{Q}_m^* = (\widehat{Q}_{p_1,m}^*, \dots, \widehat{Q}_{p_d,m}^*)$ for $p_1, \dots, p_d \in (0, 1)$. Then, we have bootstrap consistency, i.e.

$$d_{KS}(\widehat{Q}_m^*, \widehat{Q}) := \sup_{x \in \mathbb{R}^d} |P^*(\widehat{Q}_m^* \leq x) - P(\widehat{Q} \leq x)| \xrightarrow{P} 0.$$

Here, the short-hand $\underline{x} \leq \underline{y}$ for $\underline{x}, \underline{y} \in \mathbb{R}^d$ is used to denote $x_i \leq y_i$ for all $i = 1, \dots, d$.

To capture the dependence structure of the process $(X_t)_{t \in \mathbb{Z}}$ in the time series case, we approach an m -out-of- n (moving) block bootstrap procedure:

- Step 1. Choose a bootstrap sample size m , a block length l and let $b = \lceil m/l \rceil$ be the smallest number of blocks required to get a bootstrap sample of length $bl \geq m$. Define blocks $B_{i,l} = (X_{i+1}, \dots, X_{i+l})$, $i = 0, \dots, n-l$ and let i_0, \dots, i_{b-1} be i.i.d. random variables uniformly distributed on the set $\{0, 1, 2, \dots, n-l\}$.
- Step 2. Lay the blocks $B_{i_0,l}, \dots, B_{i_{b-1},l}$ end-to-end together to get

$$\begin{aligned} B_{i_0,l}, \dots, B_{i_{b-1},l} &= X_{i_0+1}, \dots, X_{i_0+l}, X_{i_1+1}, \dots, X_{i_1+l}, \dots, X_{i_{b-1}+1}, \dots, X_{i_{b-1}+l} \\ &= X_1^*, \dots, X_{bl}^* \end{aligned}$$

and discard the last $bl - m$ values to get a bootstrap sample X_1^*, \dots, X_m^* .

An application of this block bootstrap is in particular necessary to obtain bootstrap consistency if several quantiles are considered jointly. This leads to the following theorem.

Theorem 9 (Consistency of the block-wise m -out-of- n bootstrap) *Let X_1, \dots, X_n be discretely distributed random variables with $\text{supp}(P^{X_1}) = V$ as above that are observations of a strictly stationary and τ -dependent process $(X_t)_{t \in \mathbb{Z}}$ with $\sum_{h=0}^\infty h \tau(h) < \infty$. We apply the block-wise m -out-of- n bootstrap to get a bootstrap sample X_1^*, \dots, X_m^* . Suppose that $m/n + l/m + 1/l = o(1)$ as $n \rightarrow \infty$. With the notation of Theorem 8, we have bootstrap consistency, i.e.*

$$d_{KS}(\widehat{Q}_m^*, \widehat{Q}) \xrightarrow{P} 0.$$

Remark 10 It can be seen from Theorem 6(iii) that $P(\widehat{Q}_p = Q_p) \rightarrow P(Z \geq 0) = 1/2$ as $n \rightarrow \infty$ if $F(Q_p) = p$. Here, Z is a centered normal variable whose variance depends on the dependence structure of the underlying process. However, for the limit behavior of the sample quantile itself the variance of Z is not relevant and we only require symmetry around the origin. In the case of $F(Q_p) > p$ the proof of $P(\widehat{Q}_p = Q_p) \rightarrow 1$ is based on the WLLN which holds for i.i.d. as well as for τ -weakly dependent data. This implies in particular that to mimic the asymptotic behavior of a *single* quantile correctly we do not have to imitate the dependence structure correctly. Hence, the i.i.d. m -out-of- n -bootstrap is also valid for sequences of weakly dependent random variables if single quantiles are considered; for details follow the lines of the proof of Theorem 8. A similar phenomenon occurs when m -out-of- n bootstrap is used to mimic the distribution of $M_n = \max(X_1, \dots, X_n)$; see Theorem 4 and Section 4 in Athreya et al. (1999).

2.3 Randomized construction of confidence sets

In discrete setups it is more appropriate to work with confidence sets rather than confidence intervals for population quantiles. By consistency of the non-centered m -out-of- n i.i.d. bootstrap (and the m -out-of- n block bootstrap) we can apply this method to derive such confidence sets. Due to the discreteness of the underlying distribution a naive construction of confidence sets will be too conservative, that is, the effective limiting coverage of an asymptotic $(1 - \alpha)$ -quantile is strictly larger than $1 - \alpha$; actually equal to one if $\alpha < 1/2$. If one does not want to use conservative confidence sets with (too) large coverages, one can compensate this effect by randomization techniques. More precisely, we proceed as follows: We calculate one confidence set for the *sample quantile* with coverage larger than the prescribed size $1 - \alpha$ and another one with a coverage (asymptotically) smaller than $1 - \alpha$. Then, we choose randomly (with an appropriate distribution) one of these sets and use this to construct a final confidence set for the population quantile of asymptotic level $1 - \alpha$. Another difficulty that has to be taken into account is that we have bootstrap consistency only without centering, that is,

$$P^*(\widehat{Q}_{p,m}^* \leq x) \approx P(\widehat{Q}_p \leq x), \text{ but } P^*(\widehat{Q}_{p,m}^* - \widehat{Q}_p \leq x) \not\approx P(\widehat{Q}_p - Q_p \leq x), \tag{21}$$

such that the standard construction of bootstrap confidence intervals is not possible. Let V_n denote the support of the empirical marginal distribution based on X_1, \dots, X_n . Then, we define large and small confidence sets CS_L and CS_S , respectively, for the sample quantile

$$CS_L = \left[F_{\widehat{Q}_{p,m}^*}^{*-1}(\alpha/2), F_{\widehat{Q}_{p,m}^*}^{*-1}(1 - \alpha/2) \right] \cap V_n,$$

$$CS_S = \left[F_{\widehat{Q}_{p,m}^*}^{*-1}(\alpha/2), F_{\widehat{Q}_{p,m}^*}^{*-1}(1 - \alpha/2) \right) \cap V_n,$$

and their coverages

$$cov_L = P^*(\widehat{Q}_{p,m}^* \in CS_L), \quad cov_S = P^*(\widehat{Q}_{p,m}^* \in CS_S).$$

Note that $cov_L \geq 1 - \alpha$ while the size of cov_S is not clear in finite samples. It will turn out to be less than $1 - \alpha$ in the limit. Finally, we specify

$$p^* = \frac{1 - \alpha - cov_S}{cov_L - cov_S}$$

and define the bootstrap approximation of the confidence set for the sample quantile

$$\widetilde{CS} = \begin{cases} CS_L & \text{if } Y \leq p^* \\ CS_S & \text{if } Y > p^*, \end{cases}$$

where $Y \sim \text{Unif}(0, 1)$ is chosen independently from all observations and all bootstrap variables. A corresponding confidence set for the population quantile is then given by

$$CS = \widetilde{CS} - \widehat{Q}_p + H(\widehat{Q}_{p,m}^*).$$

Due to (21) and as $P(\widehat{Q}_p \in \widetilde{CS}) \rightarrow 1 - \alpha$ holds, the use of a correction term $H(\widehat{Q}_{p,m}^*) := F_{\widehat{Q}_{p,m}^*}^{*-1}(0.4)$ is necessary as an approximation of the true quantile Q_p ; see proof of Theorem 11 below. In principle, any value in $(0, 1/2]$ can be used instead of 0.4.

Theorem 11 *Suppose that either the assumptions of Theorem 8 or Theorem 9 hold true. Then, for $\alpha \in (0, 1/2)$, we have*

$$P(Q_p \in CS) \xrightarrow{n \rightarrow \infty} 1 - \alpha.$$

Remark 12 (On the use of V or V_n) The effect of using V or V_n is asymptotically negligible. For applications it might be reasonable to assume either that V is known in advance or that it is unknown. In the first case V should be used to construct the confidence intervals and in the latter case V_n seems to be the more reasonable choice.

2.4 Simulations

In this section, we illustrate the bootstrap performance by means of coverage rates of $(1 - \alpha)$ -confidence sets CS for $\alpha = 0.05$ as proposed in the previous section. To cover both cases of i.i.d. as well as time series data, let X_1, \dots, X_n be either

(a) an i.i.d. realization of a binomial distribution $X_i \sim \text{Bin}(N, \theta)$

or

(b) a realization of a (Poisson-)INAR(1) model $X_t = \beta \circ X_{t-1} + \epsilon_t$, where $\epsilon_t \sim \text{Poi}(\eta(1 - \beta))$ is Poisson distributed and $\beta \circ k \sim \text{Bin}(k, \beta)$ for $k \in \mathbb{N}_0$ denotes the binomial thinning operator.

The quantity of interest is the (sample) median, where we consider different parameter settings for both cases (a) and (b) that lead to degenerate one-point as well as non-degenerate two-point limiting distributions, respectively. In all simulations we have used V to construct confidence sets; compare Remark 12.

In Table 1, we show coverage rates of confidence sets for model a) for several sample sizes $n \in \{100, 500, 1000, 5000\}$ and parameter settings with $\theta = 0.5$ and $N \in \{1, 2, 19, 20, 39, 40\}$, where odd N leads to a non-degenerate limiting distribution ($N = 1$ is the fair coin flip) and even N results in a degenerate one-point limiting distribution. In Fig. 3, we show typical bootstrap confidence sets for the examples $\text{Bin}(19, 0.5)$ and $\text{Bin}(39, 0.5)$. As our theory provided in Sect. 2.2 suggests, we use the m -out-of- n bootstrap to mimic correctly the limiting behavior of sample quantiles in the degenerate as well as the non-degenerate case. To illustrate how sensitive the

Table 1 Coverage rates of $(1 - \alpha)$ -bootstrap confidence sets CS with $\alpha = 0.05$ for the median X_{med} of $X_t \sim Bin(N, 0.5)$ for several choices of N , sample sizes n and bootstrap sample sizes m

i.i.d.		$N = 1$				$N = 2$			
m	n	100	500	1000	5000	100	500	1000	5000
$n^{1/2}$		0.955	0.947	0.947	0.951	0.975	0.963	0.953	0.941
$n^{2/3}$		0.964	0.983	0.973	0.957	0.962	0.957	0.952	0.941
$n^{3/4}$		0.980	0.973	0.978	0.982	0.954	0.946	0.955	0.95
CLT approx.		0.548	0.502	0.507	0.503	1.000	1.000	1.000	1.000
i.i.d.		$N = 19$				$N = 20$			
m	n	100	500	1000	5000	100	500	1000	5000
$n^{1/2}$		0.995	0.994	0.976	0.952	0.894	0.977	1.000	1.000
$n^{2/3}$		0.993	0.986	0.973	0.957	0.880	0.993	0.984	0.945
$n^{3/4}$		0.988	0.973	0.978	0.982	0.909	0.974	0.975	0.95
CLT approx.		0.547	0.502	0.507	0.503	0.916	1.000	1.000	1.000
i.i.d.		$N = 39$				$N = 40$			
m	n	100	500	1000	5000	100	500	1000	5000
$n^{1/2}$		0.996	1.000	1.000	0.980	0.980	0.942	0.980	1.000
$n^{2/3}$		0.989	0.995	0.977	0.957	0.939	0.988	0.998	0.974
$n^{3/4}$		0.989	0.979	0.978	0.982	0.910	0.976	0.986	0.953
CLT approx.		0.542	0.502	0.507	0.503	0.777	0.997	1.000	1.000

bootstrap reacts on the choice of the bootstrap sample size, we show results for several (rounded) values of $m \in \{n^{1/2}, n^{2/3}, n^{3/4}\}$. For each parameter setting, we generate $K = 1000$ time series and $B = 1000$ bootstrap replicates are used to construct the confidence set as described in Sect. 2.3. By the de Moivre–Laplace theorem, a binomial probability mass function can be well approximated by a normal density around the mean if the number of experiments N is large. As suggested by one referee we included also coverage rates of the corresponding asymptotic confidence intervals. To construct these intervals, we treated the data as being normally distributed with mean Np and variance $Np(1 - p)$ such that a Bahadur-type CLT

$$\sqrt{n}(\widehat{Q}_p - Q_p) = \frac{1}{f(Q_p)} \sqrt{n}(p - \widehat{F}_n(Q_p)) + o_P(1) \xrightarrow{D} \mathcal{N}(0, S), \quad (22)$$

would hold true with limiting variance $S = \text{var}(1(X_1 \leq Q_p))/f^2(Q_p)$ if X_1, \dots, X_n are i.i.d. and $f(\cdot) = \varphi((\cdot - N\theta)/\sqrt{N\theta(1 - \theta)})/\sqrt{N\theta(1 - \theta)}$, where φ denotes the probability density function of the standard normal distribution. Precisely, we used the confidence intervals $[\widehat{X}_{med} - q_{1-\alpha/2} * \widehat{S}/\sqrt{n}, \widehat{X}_{med} - q_{\alpha/2} * \widehat{S}/\sqrt{n}]$, where \widehat{S} is an empirical version of S and q_α denotes the α -quantile of the standard normal distribution. Table 1 reports a good overall finite sample performance of our procedure,

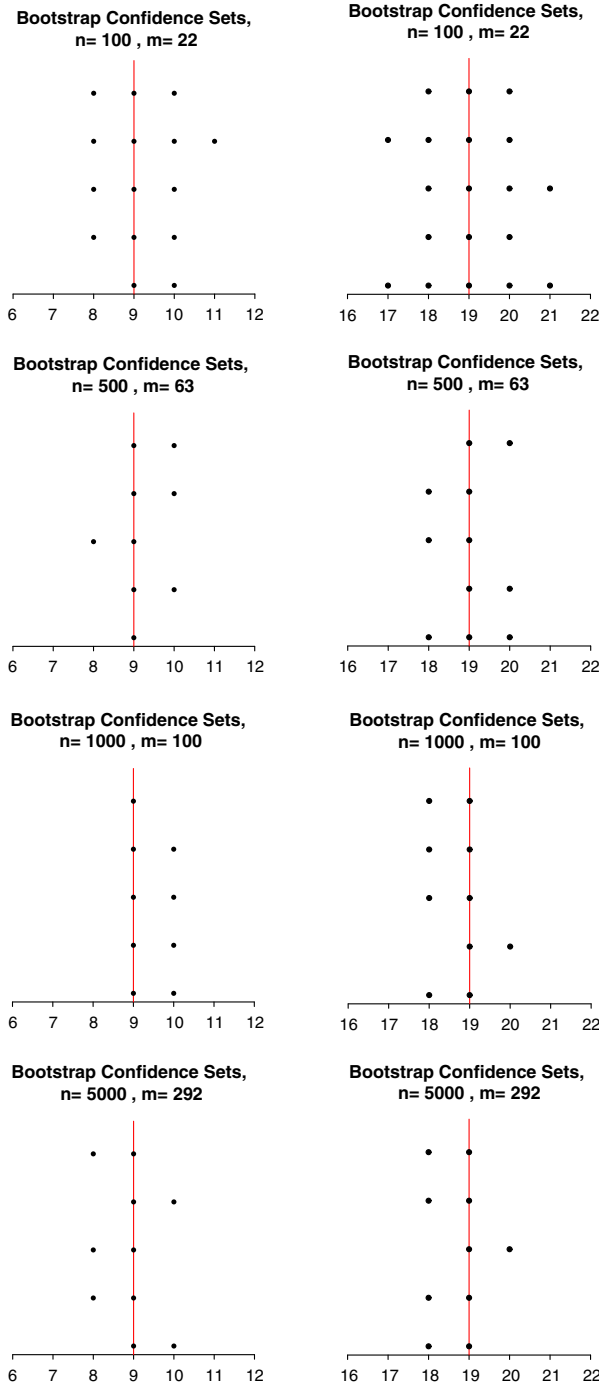


Fig. 3 Confidence sets CS for the median X_{med} from five realizations of X_1, \dots, X_n with $X_i \sim \text{Bin}(N, 0.5)$ i.i.d. for $N = 19$ (left panels) and $N = 39$ (right panels), several sample sizes n and bootstrap sample sizes $m = n^{2/3}$. The true median is marked with a red vertical line

whereas the coverage rates of the CLT-based confidence intervals that treat the data as being normally distributed clearly fail here. In the case of a limiting two-point distribution, the coverage is around 50% and in the limiting one-point distribution it converges to 100%. Observe that these confidence intervals are just ad hoc and not asymptotically valid such that we did not expect good results here. For our procedure, we see that an increasing binomial parameter N leads to higher variance of the data generating process, i.e. $\text{var}(X_t) = N/4$. Hence, confidence sets are larger and we observe a slight overcoverage. Moreover, we observe that confidence sets for even N are more conservative than for odd N which is due to the degeneracy of the limit distribution of the sample median for even N .

Remark 13 Theorem 11 yields asymptotic validity of the procedure under the minimal condition $m = o(n)$ on the intensity parameter. The question of its optimal choice has been discussed in different contexts. We refer the reader to an overview on the literature in Santana (2009, Sections 4.4 and 5.5) and Bickel and Sakov (2008). Unfortunately, these findings cannot be applied directly in the present context since their prerequisites are violated in our special case. Bickel and Sakov (2008) require the limiting distribution of the quantity under consideration to be non-degenerate which contradicts our Theorem 6(i). However, our simulation study shows that the bootstrap method is robust for different choices of the intensity m . If N is large, small choices of m lead to more conservative confidence intervals than large ones. The effect of overcoverage can be explained by larger variability caused by small bootstrap sample sizes m . It seems that $m = n^{2/3}$ is a good compromise and might be recommended as a rule of thumb. A detailed investigation of the optimal choice of m is left to further research.

In the setup (b), displayed in Table 2, we consider again the non-degenerate case for $\lambda = 3.67206\dots$ such that $X_{\text{med}} = 3$ as well as the degenerate case for $\lambda = 4$ such that $X_{\text{med}} = 4$. As discussed in Remark 10, Table 2 shows that already the i.i.d. low-intensity bootstrap leads to valid results and the block bootstrap does not lead to visible improvements of the performance.

3 Mid-distribution quantiles

3.1 Asymptotics for sample mid-quantiles

Suppose we observe X_1, \dots, X_n from a (τ -dependent) process with discrete support $\text{supp}(P^{X_1}) = V$ as defined in (19). Instead of considering classical quantiles as in Sect. 2 of the present paper, Parzen (1997, 2004) and Ma et al. (2011) suggested to investigate a modified quantile function of the corresponding so-called mid-distribution function F_{mid} , which is given by

$$F_{\text{mid}}(x) = F(x) - 0.5p(x), \quad x \in \mathbb{R},$$

where, as before, F denotes the cdf of the random variable X with probability mass function $p(x) = P(X = x)$. Their concept allows for a meaningful interpretation of

Table 2 Coverage rates of $(1 - \alpha)$ -bootstrap confidence sets CS with $\alpha = 0.05$ for the median X_{med} of the INAR(1) model $X_t = \beta \circ X_{t-1} + \epsilon_t$, $\beta = 0.5$ for two choices of λ and several sample sizes n and bootstrap sample sizes m

INAR		$\eta = 3.67206 \dots$, i.i.d.				$\eta = 4$, i.i.d.			
m	n	100	500	1000	5000	100	500	1000	5000
	$n^{1/2}$	0.989	0.993	0.963	0.972	0.801	0.778	0.882	0.957
	$n^{2/3}$	0.978	0.989	0.977	0.969	0.800	0.923	0.943	0.962
	$n^{3/4}$	0.985	0.980	0.984	0.988	0.820	0.901	0.940	0.939
INAR		$\eta = 3.67206 \dots$, MBB, $l = m^{1/2}$				$\eta = 4$, MBB, $l = m^{1/2}$			
m	n	100	500	1000	5000	100	500	1000	5000
	$n^{1/2}$	0.989	0.990	0.969	0.966	0.829	0.762	0.868	0.962
	$n^{2/3}$	0.986	0.982	0.971	0.980	0.805	0.930	0.928	0.953
	$n^{3/4}$	0.981	0.979	0.986	0.986	0.818	0.895	0.942	0.955

Results for i.i.d. resampling (i.i.d., upper tables) and the moving block bootstrap (MBB, lower tables) with block length $l = m^{1/2}$ are given

quantiles in the discrete setup and appears to be beneficial in cases of tied samples. Here, we refer to the paper of [Ma et al. \(2011\)](#) for details. In particular, it is argued there that mid-quantiles, related to the mid-distribution function in the sense of (23) below, behave more favorably than classical quantiles. That is, contrary to classical sample quantiles in discrete setups, they showed that sample (mid-)quantiles based on the mid-distribution function converge to non-degenerate limiting distributions when properly centered and inflated with the usual \sqrt{n} -rate as long as they do not correspond to the boundary values of the support of the underlying distribution. In the latter case the limiting distribution is degenerate for any choice of the inflation factor. Moreover, they show that asymptotic theory coincides for mid-quantiles and classical quantiles if the underlying distribution is absolutely continuous. In view of this, mid-quantiles can be interpreted as a natural generalization of classical quantiles which appears to be robust to discreteness of the underlying distribution.

We first assume the support V to be bounded, $V = \{v_1, \dots, v_d \mid v_1 < \dots < v_d\}$, say. However, it turns out that the case of unbounded support can be treated similarly and the asymptotics are even easier; see Remark 15 below. According to [Ma et al. \(2011\)](#) the mid-quantile function is a linear interpolation of the points $(F_{mid}(v_j), v_j)$, $j = 1, \dots, d$. More precisely, we define the p th population mid-quantile $Q_{p, mid}$ as

$$Q_{p, mid} = \begin{cases} v_1 & \text{if } p < F_{mid}(v_1) \\ v_k & \text{if } p = F_{mid}(v_k), k = 1, \dots, d \\ \lambda v_k + (1 - \lambda)v_{k+1} & \text{if } p = \lambda F_{mid}(v_k) + (1 - \lambda)F_{mid}(v_{k+1}), \lambda \in (0, 1), \\ & k = 1, \dots, d - 1 \\ v_d & \text{if } p > F_{mid}(v_d) \end{cases} \quad (23)$$

and its empirical counterpart $\widehat{Q}_{p, mid}$ as

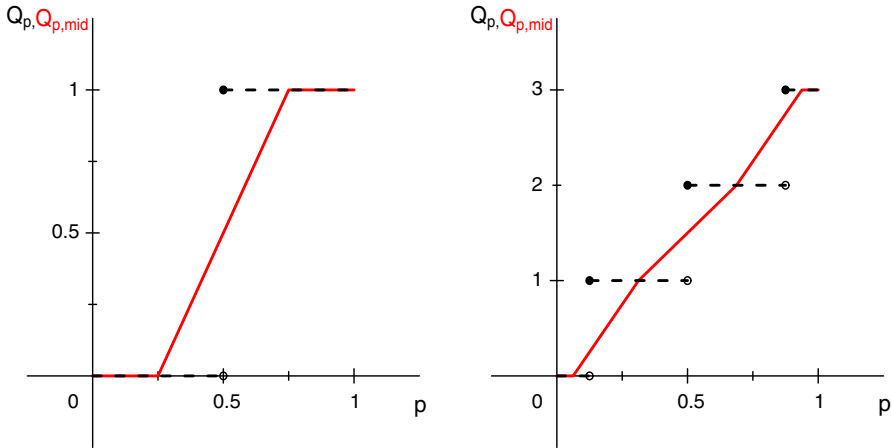


Fig. 4 Comparison of quantile function (black, dashed) and mid-quantile function (red, solid) for Bin(1, 0.5) (left panel) and Bin(3, 0.5) (right panel)

$$\widehat{Q}_{p, \text{mid}} = \begin{cases} v_1 & \text{if } p < \widehat{F}_{\text{mid}}(v_1) \\ v_k & \text{if } p = \widehat{F}_{\text{mid}}(v_k) < \widehat{F}_{\text{mid}}(v_{k+1}), k = 1, \dots, d \\ \lambda_n v_k + (1 - \lambda_n)v_{k+1} & \text{if } p = \lambda \widehat{F}_{\text{mid}}(v_k) + (1 - \lambda_n)\widehat{F}_{\text{mid}}(v_{k+1}), \lambda_n \in (0, 1), \\ & \widehat{F}_{\text{mid}}(v_k) < \widehat{F}_{\text{mid}}(v_{k+1}), k = 1, \dots, d - 1 \\ v_d & \text{if } p > \widehat{F}_{\text{mid}}(v_d) \end{cases} \quad (24)$$

where $\widehat{F}_{\text{mid}}(x) = n^{-1} \sum_{k=1}^n \{1(X_k \leq x) - 0.5 \cdot 1(X_k = x)\}$ is the empirical counterpart of $F_{\text{mid}}(x)$; see also Fig. 4 for illustration. There, we compare the classical and the mid-quantile function for the Bin(1, 0.5) and the Bin(3, 0.5) distribution. For both distributions the classical median cannot be estimated consistently which follows from the plateau of the classical quantile function from argument 0.5 onwards. Contrary, the mid-quantile function is strictly increasing around this argument which then results in a \sqrt{n} -consistency of its empirical version, see also Theorem 14 below.

Our first goal is to extend the asymptotic results of Ma et al. (2011) from i.i.d. data to strictly stationary, τ -dependent processes. Similar to Sect. 2 of the paper, any other concept of dependence might be applied as long as the CLT for the empirical distribution function holds. For sake of definiteness, we restrict ourselves to τ -dependence here.

Theorem 14 (Asymptotics of sample mid-quantiles for discrete distributions) *Suppose that X_1, \dots, X_n are either i.i.d. or observations of a strictly stationary, τ -dependent process $(X_t)_{t \in \mathbb{Z}}$ with $\sum_{h=0}^{\infty} \tau(h) < \infty$. Let the support of P^{X_1} be $V = \{v_1, \dots, v_d\}$ such that $v_1 < \dots < v_d$ and denote the corresponding probabilities by a_1, \dots, a_d . Further, define $a_0 = a_{d+1} = 0$, $v_0 = v_1$ and $v_{d+1} = v_d$. Then, we have*

$$\sqrt{n}(\widehat{Q}_{p,\text{mid}} - Q_{p,\text{mid}}) \xrightarrow{\mathcal{D}} \begin{cases} 0 & \text{if } p < F_{\text{mid}}(v_1) \text{ or } p > F_{\text{mid}}(v_d) \\ Z_1 & \text{if } p = \lambda F_{\text{mid}}(v_{k+1}) + (1 - \lambda)F_{\text{mid}}(v_{k+2}), \lambda \in (0, 1), \\ & k = 0, \dots, d - 2 \\ Z_2 & \text{if } p = F_{\text{mid}}(v_{k+1}), k = 1, \dots, d - 2 \\ Z_3 & \text{if } p = F_{\text{mid}}(v_1) \\ Z_4 & \text{if } p = F_{\text{mid}}(v_d) \end{cases} \tag{25}$$

where Z_1, Z_2, Z_3, Z_4 are random variables having certain non-degenerate distributions as described in the following. Z_1 is centered and normally distributed with variance

$$\sigma_1^2 = 4 \left(\frac{v_{k+1} - v_{k+2}}{a_{k+2} + a_{k+1}} \right)^2 h'_{k+2} \Sigma^{(k+2)} h_{k+2}, \tag{26}$$

where

$$h_{k+2} = \left(1, \dots, 1, 1 - \frac{F_{\text{mid}}(v_{k+2}) - p}{a_{k+1} + a_{k+2}}, \frac{1}{2} - \frac{F_{\text{mid}}(v_{k+2}) - p}{a_{k+1} + a_{k+2}} \right)'$$

and $\Sigma^{(k+2)} = (\Sigma_{j_1, j_2})_{j_1, j_2=1, \dots, k+2}$ with $\Sigma_{j_1, j_2} = \sum_{h \in \mathbb{Z}} \text{cov}(1(X_h = v_{j_1}), 1(X_0 = v_{j_2}))$. The density of Z_2 is that of a centered normal distribution with variance

$$\sigma_{2-}^2 = 4 \left(\frac{v_{k+1} - v_k}{a_k + a_{k+1}} \right)^2 \left\{ (1, \dots, 1, 0.5) \Sigma^{(k+1)} (1, \dots, 1, 0.5)' \right\}$$

on the negative real line and that of a centered normal distribution with variance

$$\sigma_{2+}^2 = 4 \left(\frac{v_{k+2} - v_{k+1}}{a_{k+1} + a_{k+2}} \right)^2 \left\{ (1, \dots, 1, 0.5) \Sigma^{(k+1)} (1, \dots, 1, 0.5)' \right\}$$

on the positive real line; such distributions are termed half-Gaussian or two-piece normal distributions. The distribution of Z_3 has point mass of 1/2 in zero and admits a density on the positive real line which is that of a centered normal distribution with variance σ_{2+}^2 . Similarly, Z_4 has point mass of 1/2 in zero and admits a density on the negative real line which is that of a centered normal distribution with variance σ_{2-}^2 .

Observe that depending on the situation, the limiting results established in Theorem 14 include four different types of distributions. These are, degenerate, Gaussian, half-Gaussian and half-Gaussian with point masses at the boundary. Also observe that we present the limiting results for sample mid-quantiles in a different way than Ma et al. (2011). The results displayed in (25) will turn out to be convenient for investigating the applicability of bootstrap methods in the sequel. Nevertheless, in comparison to the i.i.d. case, only the covariance matrix $\Sigma^{(k+2)}$ changes.

Remark 15 (Boundary issues)

- (i) In the boundary cases $p < F_{\text{mid}}(v_1)$ and $p > F_{\text{mid}}(v_d)$ we even get $\widehat{Q}_{p,\text{mid}} = Q_{p,\text{mid}}$ with probability tending to one; see the proof of Theorem 14. These stronger results are used in the proofs of bootstrap consistency later on.
- (ii) Note that the results of Theorem 14 carry over to countable support V as long as it does not contain an accumulation point. Then, the cases $p < F_{\text{mid}}(v_1)$ and/or $p > F_{\text{mid}}(v_d)$ simply disappear; see also Remark 2 in Ma et al. (2011).

Remark 16 Similar to Theorem 6, it is possible to prove joint convergence of several sample mid-quantiles. For clarity of exposition, we do not give the exact convergence results here, but mention that multivariate limiting distributions of several sample mid-quantiles can be obtained essentially by combining the univariate results of Theorem 14 above.

Before considering the bootstrap for mid-quantiles in Sect. 3.2, we first illustrate the concept of mid-quantiles with the help of a continuation of the coin flip example discussed in the Introduction; compare also Fig. 4.

Toy example: coin flip data for mid-quantiles Suppose a fair coin is flipped independently n times and we observe a sequence X_1, \dots, X_n of zeros and ones such that $P(X_t = 0) = 1/2 = 1 - P(X_t = 1)$. Let $X_{\text{med, mid}} = Q_{0.5,\text{mid}}$ and $\widehat{X}_{\text{med, mid}} = \widehat{Q}_{0.5,\text{mid}}$ denote the population mid-median and the sample mid-median, respectively. Then, (23) gives $X_{\text{med, mid}} = 1/2$ and from Theorem 14, we get

$$\sqrt{n}(\widehat{X}_{\text{med, mid}} - X_{\text{med, mid}}) \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1/4). \tag{27}$$

Thus, the sample mid-median fulfils a CLT and, in particular, it is a \sqrt{n} -consistent estimator for the mid-median.

3.2 Bootstrapping sample mid-quantiles

We showed that standard bootstrap proposals may fail in the purely discrete data case for classical sample quantiles. A closer inspection of the bootstrap invalidity result of Theorem 1 shows that this issue is caused by the discreteness of the distributions which in turn leads to quantile functions having jumps. In view of this observation, the use of mid-quantiles may circumvent this problem, because the corresponding mid-quantile function is piecewise linear and thus, in particular, continuous by construction; compare Fig. 4.

In a first step, we investigate to what extend m -out-of- n -type bootstraps (i.i.d. and block) are capable to mimic correctly the limiting distributions established in Theorem 14. Here, we allow explicitly the case $m = n$ to cover also Efron’s bootstrap and the standard moving block bootstrap. To fix some notation, let $\widehat{Q}_{p,\text{mid},m}^*$ denote the p th bootstrap sample mid-quantile based on bootstrap observations X_1^*, \dots, X_m^* . More precisely and analogue to (24), we define

$$\widehat{Q}_{p,\text{mid},m}^* = \begin{cases} v_1 & \text{if } p < \widehat{F}_{\text{mid},m}^*(v_1) \\ v_k & \text{if } p = \widehat{F}_{\text{mid},m}^*(v_k) < \widehat{F}_{\text{mid},m}^*(v_{k+1}), k = 1, \dots, d \\ \lambda_m^* v_k + (1 - \lambda_m^*) v_{k+1} & \text{if } p = \lambda_m^* \widehat{F}_{\text{mid},m}^*(v_k) + (1 - \lambda_m^*) \widehat{F}_{\text{mid},m}^*(v_{k+1}), \lambda_m^* \in (0, 1), \\ & \widehat{F}_{\text{mid},m}^*(v_k) < \widehat{F}_{\text{mid},m}^*(v_{k+1}), k = 1, \dots, d - 1 \\ v_d & \text{if } p > \widehat{F}_{\text{mid},m}^*(v_d) \end{cases} \quad (28)$$

where $\widehat{F}_{\text{mid},m}^*(x) = m^{-1} \sum_{k=1}^m \{1(X_k^* \leq x) - 0.5 \cdot 1(X_k^* = x)\}$ is the bootstrap counterpart of $\widehat{F}_{\text{mid}}(x)$ based on X_1^*, \dots, X_m^* .

Theorem 17 (Asymptotics of bootstrap sample mid-quantiles for discrete distributions) *Suppose either (i) or (ii) holds, where*

- (i) X_1, \dots, X_n are i.i.d. and we draw i.i.d. bootstrap replicates X_1^*, \dots, X_m^* such that $m \rightarrow \infty$ and $m = o(n)$ or $m = n$ as $n \rightarrow \infty$
- (ii) X_1, \dots, X_n are τ -dependent with $\sum_{h=1}^\infty h\tau(h) < \infty$ and we apply an m -out-of- n block bootstrap with block length l to get X_1^*, \dots, X_m^* such that $l/m + 1/l = o(1)$ and $m = o(n)$ or $m = n$ as $n \rightarrow \infty$

Then, we have

$$\sqrt{m}(\widehat{Q}_{p,\text{mid},m}^* - \widehat{Q}_{p,\text{mid}}) \xrightarrow{\mathcal{D}} \begin{cases} 0 & \text{if } p < F_{\text{mid}}(v_1) \text{ or } p > F_{\text{mid}}(v_d) \\ Z_1 & \text{if } p = \lambda F_{\text{mid}}(v_{k+1}) + (1 - \lambda)F_{\text{mid}}(v_{k+2}), \lambda \in (0, 1), \\ & k = 0, \dots, d - 2 \end{cases} \quad (29)$$

and

$$\sqrt{m}(\widehat{Q}_{p,\text{mid},m}^* - Q_{p,\text{mid}}) \xrightarrow{\mathcal{D}} \begin{cases} Z_2 & \text{if } p = F_{\text{mid}}(v_{k+1}), k = 1, \dots, d - 2 \\ Z_3 & \text{if } p = F_{\text{mid}}(v_1) \\ Z_4 & \text{if } p = F_{\text{mid}}(v_d) \end{cases} \quad (30)$$

in probability, respectively. The distributions of Z_1 to Z_4 are described in Theorem 14.

At this point, it is worth noting that the results of Theorem 17 above do not require at all the use of an m -out-of- n -type bootstrap procedure with $m = o(n)$ to mimic correctly the complicated limiting distributions in all cases presented in Theorem 14. However, a comparison of (25) with (29) and (30) shows that the correct centering for the bootstrap sample mid-quantiles depends on the true situation. That is, $\widehat{Q}_{p,\text{mid},m}^*$ has to be centered around the sample mid-quantile $\widehat{Q}_{p,\text{mid}}$ for the first two cases and around the population quantile $Q_{p,\text{mid}}$ for the latter three. However, as the true mid-quantile function is generally unknown, the true situation is also not known. Consequently, the results of Theorem 17 are per se useless for practical applications as it is not clear which centering has to be used.

To overcome this issue, we require the bootstrap procedure to be valid for all different cases when centered around one and the same quantity. To achieve this, note that the difference of the left-hand sides of (29) and (30) computes to

$$\begin{aligned} \sqrt{m}(\widehat{Q}_{p,\text{mid},m}^* - Q_{p,\text{mid}}) - \sqrt{m}(\widehat{Q}_{p,\text{mid},m}^* - \widehat{Q}_{p,\text{mid}}) &= \sqrt{\frac{m}{n}} \{ \sqrt{n}(\widehat{Q}_{p,\text{mid}} - Q_{p,\text{mid}}) \} \\ &= O_P \left(\sqrt{\frac{m}{n}} \right) \end{aligned} \tag{31}$$

and vanishes for $m = o(n)$, but not for $m = n$. This leads to the following result.

Corollary 18 (Consistency of m -out-of- n bootstraps for sample mid-quantiles) *Suppose either (i) or (ii) in Theorem 17 holds with $m = o(n)$. Then, we have*

$$d_{\text{KS}} \left(\sqrt{m}(\widehat{Q}_{p,\text{mid},m}^* - \widehat{Q}_{p,\text{mid}}), \sqrt{n}(\widehat{Q}_{p,\text{mid}} - Q_{p,\text{mid}}) \right) \xrightarrow{\mathcal{P}} 0.$$

3.3 Randomized construction of confidence intervals

We invoke the ideas of Sect. 2.3 to construct confidence intervals of level $1 - \alpha$ for mid-quantiles. These quantities take their values in the interval $[v_1, v_d]$ if $V = \{v_1, \dots, v_d\}$ such that $v_1 < \dots < v_d$ in contrast to classical quantiles that take their values only in the countable set V . In particular, if the image of the mid-quantile function is the whole real line, the limit distribution is continuous by Theorems 14 and 17. Therefore, no randomization techniques are required to construct asymptotic exact $(1 - \alpha)$ confidence sets. If this is not the case, a randomization procedure as described in the sequel has to be applied. Note that the asymptotics in the previous section do not rely on the (empirical) mid-quantile itself but on suitably centered and inflated versions. Therefore, instead of CS_L and CS_S defined in Sect. 2.3, we consider large and small intervals of the form

$$\begin{aligned} \text{CI}_{L,\text{mid}} &= \left[F_{\sqrt{m}(\widehat{Q}_{p,\text{mid},m}^* - \widehat{Q}_{p,\text{mid}})}^{*-1}(\alpha/2), F_{\sqrt{m}(\widehat{Q}_{p,\text{mid},m}^* - \widehat{Q}_{p,\text{mid}})}^{*-1}(1 - \alpha/2) \right], \\ \text{CI}_{S,\text{mid}}^{(r)} &= \left[F_{\sqrt{m}(\widehat{Q}_{p,\text{mid},m}^* - \widehat{Q}_{p,\text{mid}})}^{*-1}(\alpha/2), F_{\sqrt{m}(\widehat{Q}_{p,\text{mid},m}^* - \widehat{Q}_{p,\text{mid}})}^{*-1}(1 - \alpha/2) \right), \\ \text{CI}_{S,\text{mid}}^{(l)} &= \left(F_{\sqrt{m}(\widehat{Q}_{p,\text{mid},m}^* - \widehat{Q}_{p,\text{mid}})}^{*-1}(\alpha/2), F_{\sqrt{m}(\widehat{Q}_{p,\text{mid},m}^* - \widehat{Q}_{p,\text{mid}})}^{*-1}(1 - \alpha/2) \right) \end{aligned}$$

and their coverages

$$\begin{aligned} \text{cov}_{L,\text{mid}} &= P^* \left(\sqrt{m}(\widehat{Q}_{p,\text{mid},m}^* - \widehat{Q}_{p,\text{mid}}) \in \text{CI}_{L,\text{mid}} \right), \\ \text{cov}_{S,\text{mid}}^{(r)} &= P^* \left(\sqrt{m}(\widehat{Q}_{p,\text{mid},m}^* - \widehat{Q}_{p,\text{mid}}) \in \text{CI}_{S,\text{mid}}^{(r)} \right), \\ \text{cov}_{S,\text{mid}}^{(l)} &= P^* \left(\sqrt{m}(\widehat{Q}_{p,\text{mid},m}^* - \widehat{Q}_{p,\text{mid}}) \in \text{CI}_{S,\text{mid}}^{(l)} \right). \end{aligned}$$

Finally, we specify the probability for choosing the large interval

$$p_{\text{mid}}^* = \begin{cases} \frac{1-\alpha-\text{cov}_{S,\text{mid}}^{(r)}}{\text{cov}_{L,\text{mid}}-\text{cov}_{S,\text{mid}}^{(r)}}, & \text{cov}_{S,\text{mid}}^{(r)} \leq 1-\alpha \\ \frac{1-\alpha-\text{cov}_{S,\text{mid}}^{(l)}}{\text{cov}_{L,\text{mid}}-\text{cov}_{S,\text{mid}}^{(l)}}, & \text{otherwise} \end{cases}$$

and define the bootstrap approximation of the confidence set for the p -level mid-quantile

$$CI = \begin{cases} \left[\hat{Q}_{p,\text{mid}} - \frac{F^{*-1}(\hat{Q}_{p,\text{mid},m}^* - \hat{Q}_{p,\text{mid}})}{\sqrt{n}}^{(1-\alpha/2)}, \hat{Q}_{p,\text{mid}} - \frac{F^{*-1}(\hat{Q}_{p,\text{mid},m}^* - \hat{Q}_{p,\text{mid}})}{\sqrt{n}}^{(\alpha/2)} \right] & \text{if } Y \leq p_{\text{mid}}^* \\ \left(\hat{Q}_{p,\text{mid}} - \frac{F^{*-1}(\hat{Q}_{p,\text{mid},m}^* - \hat{Q}_{p,\text{mid}})}{\sqrt{n}}^{(1-\alpha/2)}, \hat{Q}_{p,\text{mid}} - \frac{F^{*-1}(\hat{Q}_{p,\text{mid},m}^* - \hat{Q}_{p,\text{mid}})}{\sqrt{n}}^{(\alpha/2)} \right) & \text{if } Y > p_{\text{mid}}^* \\ & \text{and } \text{cov}_{S,\text{mid}}^{(r)} \leq 1-\alpha \\ \left(\hat{Q}_{p,\text{mid}} - \frac{F^{*-1}(\hat{Q}_{p,\text{mid},m}^* - \hat{Q}_{p,\text{mid}})}{\sqrt{n}}^{(1-\alpha/2)}, \hat{Q}_{p,\text{mid}} - \frac{F^{*-1}(\hat{Q}_{p,\text{mid},m}^* - \hat{Q}_{p,\text{mid}})}{\sqrt{n}}^{(\alpha/2)} \right) & \text{otherwise} \end{cases}$$

where $Y \sim \text{Unif}(0, 1)$ is chosen independently from all observations and all bootstrap variables. This gives an asymptotic confidence interval of level $1 - \alpha$.

Theorem 19 *Suppose that the assumptions of Corollary 18 hold true. Then, for $\alpha \in (0, 1/2)$,*

$$P(Q_{p,\text{mid}} \in CI) \xrightarrow[n \rightarrow \infty]{} 1 - \alpha.$$

3.4 Simulations

In this section, we illustrate the bootstrap performance by means of coverage rates of $(1 - \alpha)$ -confidence intervals CI for $\alpha = 0.05$ as proposed in the previous section. To make the simulation results comparable to those obtained in Sect. 2.4, we use the same settings here. Recall that the image of mid-quantile functions is continuous which leads to confidence intervals rather than confidence sets; compare Fig. 5. Contrary to the results in setup (a) obtained for classical quantiles, all choices of the binomial parameter N lead to non-degenerate distributions for the sample mid-median. In view of Table 3, we observe that the bootstrap works equally well in both cases. As for the classical quantiles, we included coverage rates of asymptotic confidence intervals that treat the data as being normally distributed for the i.i.d. setup such that we can make use of the CLT established in Theorem 1, Case 1 of Ma et al. (2011). This result leads to the same limiting distribution as obtained in the CLT for classical quantiles in (22) and we used the same construction of confidence intervals as in Sect. 2.4. For the coverage rates of the CLT-based confidence intervals that treat the data as being normally distributed, we see in Table 3 a systematic overcoverage which is in general less pronounced for larger values of N . To explain the latter observation,

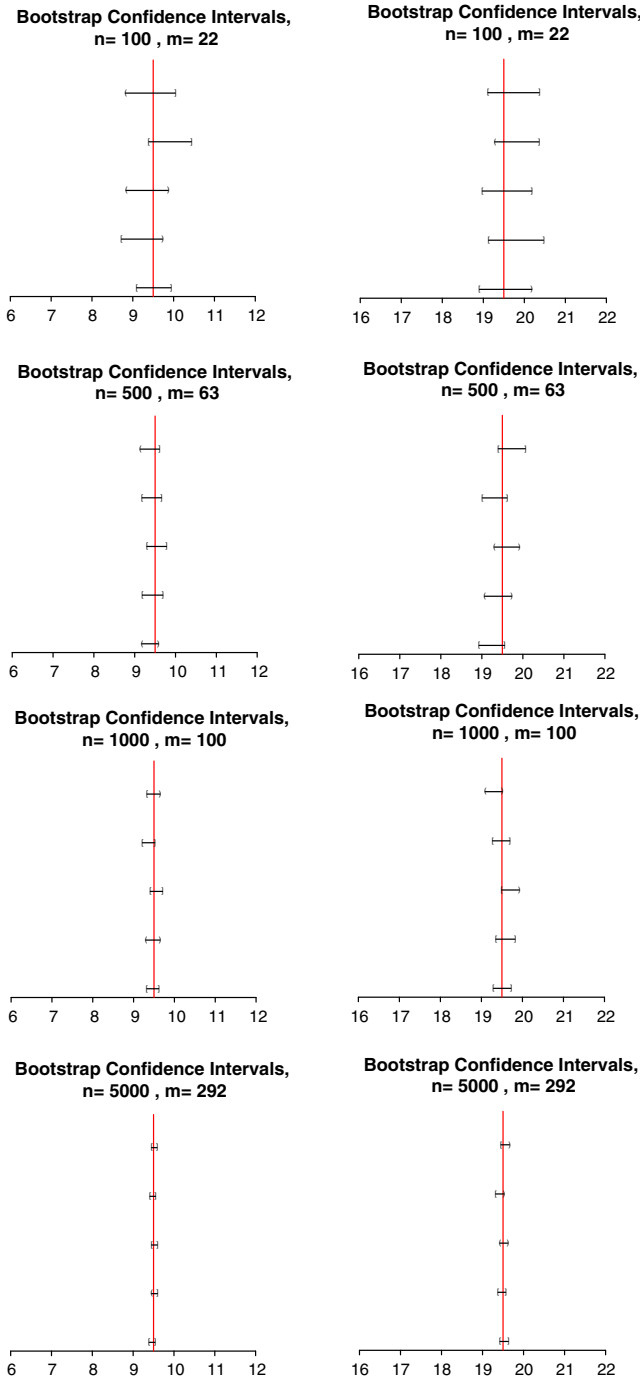


Fig. 5 Confidence intervals CI for the mid-median $X_{\text{med,mid}}$ from five realizations of X_1, \dots, X_n with $X_i \sim \text{Bin}(N, 0.5)$ i.i.d. for $N = 19$ (left panels) and $N = 39$ (right panels), several sample sizes n and bootstrap sample sizes $m = n^{2/3}$. The true mid-median is marked with a red vertical line

Table 3 Coverage rates of $(1-\alpha)$ -bootstrap confidence sets CI with $\alpha = 0.05$ for the mid-median $X_{\text{med, mid}}$ of $X_i \sim \text{Bin}(N, 0.5)$ for several choices of N , sample sizes n and bootstrap sample sizes m

iid		$N = 1$				$N = 2$			
m	n	100	500	1000	5000	100	500	1000	5000
$n^{1/2}$		0.935	0.940	0.932	0.941	0.887	0.894	0.914	0.914
$n^{2/3}$		0.927	0.942	0.935	0.943	0.902	0.909	0.925	0.928
$n^{3/4}$		0.949	0.941	0.942	0.944	0.896	0.919	0.929	0.933
CLT approx.		0.977	0.980	0.984	0.985	0.949	0.931	0.950	0.935
iid		$N = 19$				$N = 20$			
m	n	100	500	1000	5000	100	500	1000	5000
$n^{1/2}$		0.916	0.942	0.941	0.952	0.919	0.935	0.941	0.941
$n^{2/3}$		0.918	0.944	0.946	0.950	0.920	0.928	0.930	0.945
$n^{3/4}$		0.916	0.951	0.949	0.945	0.917	0.930	0.935	0.947
CLT approx.		0.956	0.965	0.971	0.966	0.962	0.965	0.973	0.961
iid		$N = 39$				$N = 40$			
m	n	100	500	1000	5000	100	500	1000	5000
$n^{1/2}$		0.906	0.943	0.941	0.949	0.922	0.932	0.943	0.956
$n^{2/3}$		0.911	0.942	0.938	0.956	0.927	0.932	0.956	0.948
$n^{3/4}$		0.909	0.941	0.938	0.954	0.932	0.939	0.945	0.943
CLT approx.		0.960	0.965	0.964	0.961	0.957	0.957	0.967	0.966

note that these confidence intervals are not asymptotically valid as they assume an underlying normal distribution. Even though the concept of mid-quantiles slightly differs from the classical ones in the discrete setup, a smooth modification of the quantile function appears to be beneficial wrt coverage rate performance of bootstrap confidence intervals.

In comparison to the results displayed in Table 2 for classical quantiles, Table 4 illustrates the necessity of a block-type resampling scheme that takes the dependence structure of the INAR process in setting (b) into account.

4 Conclusion

In this paper, we investigated bootstrap validity for classical quantiles as well as so-called mid-quantiles of discrete distributions. The classical quantile function is piecewise constant and discontinuous which makes statistical inference challenging. The concept of mid-distribution tries to overcome this deficiency by relying on piecewise linear mid-quantile functions that are continuous, but not differentiable. This approach is partly motivated by the fact that the latter function coincides with the classical quantile function if the underlying distribution is continuous. Indeed, in contrast to classical quantiles, mid-quantiles can be estimated consistently. Regarding the

Table 4 Coverage rates of $(1 - \alpha)$ -bootstrap confidence sets CI with $\alpha = 0.05$ for the mid-median $X_{\text{med, mid}}$ of the INAR(1) model $X_t = \beta \circ X_{t-1} + \epsilon_t$, $\beta = 0.5$ for two choices of λ and several sample sizes n and bootstrap sample sizes m

INAR		$\eta = 3.67206 \dots$, i.i.d.				$\eta = 4$, i.i.d.			
m	n	100	500	1000	5000	100	500	1000	5000
$n^{1/2}$		0.739	0.761	0.777	0.803	0.739	0.779	0.742	0.754
$n^{2/3}$		0.748	0.773	0.794	0.798	0.758	0.764	0.768	0.765
$n^{3/4}$		0.749	0.808	0.785	0.792	0.748	0.780	0.772	0.772

INAR		$\eta = 3.67206 \dots$, MBB, $l = m^{1/2}$				$\eta = 4$, MBB, $l = m^{1/2}$			
m	n	100	500	1000	5000	100	500	1000	5000
$n^{1/2}$		0.825	0.887	0.925	0.931	0.827	0.903	0.899	0.924
$n^{2/3}$		0.861	0.925	0.930	0.950	0.858	0.916	0.928	0.949
$n^{3/4}$		0.867	0.927	0.942	0.936	0.853	0.935	0.922	0.950

Results for i.i.d. resampling (iid, upper tables) and the moving block bootstrap (MBB, lower tables) with block length $l = m^{1/2}$ are given

Table 5 Bootstrap (in-)consistency for single sample (mid-)quantiles

Data	Method	Centering	Classical quantiles	Mid-quantiles
i.i.d.	i.i.d. bootstrap	Yes/no	X	X
	m -Out-of- n i.i.d. bootstrap	No	✓	X
	m -Out-of- n i.i.d. bootstrap	Yes	X	✓
Weakly dependent	m -Out-of- n i.i.d. bootstrap	No	✓	X
	Block bootstrap	Yes/no	X	X
	m -Out-of- n -block bootstrap	No	✓	X
	m -Out-of- n -block bootstrap	Yes	X	✓

validity of bootstrap methods this concept alone is not entirely successful. In both cases, low-intensity (block) bootstrap methods are required to mimic the distribution of the (mid-)quantile estimators correctly. In particular two tuning parameters, i.e. the intensity m and the block length l have to be chosen, irrespective of the type of quantiles. Moreover, to overcome the issue of potentially too conservative intervals, randomization techniques have to be invoked. An overview of the (in-)consistency of all bootstrap methods addressed in this paper is given in Table 5.

Still, smoothness of mid-quantile functions in comparison to ordinary quantile functions turns out to be beneficial wrt the finite sample performance. Despite the application of randomization techniques, confidence sets for classical quantiles tend to be quite conservative. This effect is not observed for the mid-distribution counterparts where bootstrap consistency for commonly centered quantities leads to a straightforward construction of confidence intervals. Therefore, the question arises whether further smooth modifications of mid-quantiles may lead to even better results. A first

attempt has been proposed by Wang and Hutson (2011) which is motivated by the Harrell–Davis quantile estimator for continuous distributions. These quantile estimators appear as sums of weighted order statistics where the weights are smooth functions of Beta cdfs. However, while Harrell and Davis (1982) use this method for the order statistic of the sample itself, Wang and Hutson (2011) apply this to the support instead. Hence, it is not clear how their definition of quantiles can be used directly for continuous data and whether there is a deep relationship between classical quantiles and these variants as in the case of mid-quantiles. Therefore, we did not follow this line of research in the present paper. Nevertheless, we conjecture that proving consistency of i.i.d. and block bootstrap methods is straightforward since the proof of asymptotic normality in Wang and Hutson (2011) relies on the CLT for the empirical cdf and the Δ -method only. The construction of other smooth modifications of quantiles and even more importantly the identification of their relationship to classical quantiles for continuous distributions and convenience for practitioners goes far beyond the scope of our paper and should be investigated in future research.

Appendix: Proofs and auxiliary results

Proofs of the main results

Proof of Theorem 1 We first prove (14). With the notation

$$Z_n = \sqrt{n} \frac{\widehat{F}_n(\epsilon) - F(\epsilon)}{\sqrt{\text{var}(1(X_1 \leq \epsilon))}} \quad \text{and} \quad Z_n^* = \sqrt{n} \frac{\widehat{F}_n^*(\epsilon) - \widehat{F}_n(\epsilon)}{\sqrt{\text{var}(1(X_1 \leq \epsilon))}}$$

for any fixed $\epsilon \in (0, 1)$ and using the fact that for any distribution function G on \mathbb{R} , $G(x) \geq t$ if and only if $x \geq G^{-1}(t)$, we get

$$\begin{aligned} P^*(\widehat{X}_{\text{med}}^* = 0) &= P^*(\widehat{X}_{\text{med}}^* \leq \epsilon) = P^*\left(\frac{1}{2} \leq \widehat{F}_n^*(\epsilon)\right) = 1 - P^*(Z_n^* < -Z_n) \\ &= 1 - \Phi(-Z_n) + (\Phi(-Z_n) - P^*(Z_n^* < -Z_n)). \end{aligned}$$

In conjunction with Polya's Theorem, we get from Lemma 21 that

$$\left| \Phi(-Z_n) - P^*(Z_n^* < -Z_n) \right| \leq \sup_{x \in \mathbb{R}} |\Phi(x) - P^*(Z_n^* < x)| = o_P(1).$$

By Slutsky's Theorem, it remains to show that

$$1 - \Phi(-Z_n) \xrightarrow{\mathcal{D}} U \sim \text{Unif}(0, 1),$$

which follows from $Z_n \xrightarrow{\mathcal{D}} Z \sim \mathcal{N}(0, 1)$, the Simulation Lemma and from $U := 1 - \widetilde{U} \sim \text{Unif}(0, 1)$ if $\widetilde{U} \sim \text{Unif}(0, 1)$. The result in (i) follows immediately from (11) and

$$|P^*(\widehat{X}_{\text{med}}^* = 0) - P(\widehat{X}_{\text{med}} = 0)| \xrightarrow{\mathcal{D}} \left|U - \frac{1}{2}\right| \sim \text{Unif}(0, 1/2).$$

Now, we show the result in (ii). As $\widehat{X}_{\text{med}}, \widehat{X}_{\text{med}}^* \in \{0, 1\}$, $X_{\text{med}} = 0$ and due to (5) and (13), we have to derive the asymptotics of the bivariate random variables

$$\begin{aligned} & \left(\begin{array}{l} P^*(\widehat{X}_{\text{med}}^* - \widehat{X}_{\text{med}} \leq -1) - P(\widehat{X}_{\text{med}} - X_{\text{med}} \leq -1) \\ P^*(\widehat{X}_{\text{med}}^* - \widehat{X}_{\text{med}} \leq 0) - P(\widehat{X}_{\text{med}} - X_{\text{med}} \leq 0) \end{array} \right) \\ &= \left(\begin{array}{l} P^*(\widehat{X}_{\text{med}}^* - \widehat{X}_{\text{med}} \leq -1) \\ P^*(\widehat{X}_{\text{med}}^* - \widehat{X}_{\text{med}} \leq 0) - \frac{1}{2} \end{array} \right) + o_P(1) \end{aligned}$$

to compute the supremum of both components. By straightforward calculations and due to $P^*(\widehat{X}_{\text{med}}^* \leq 0) = 1 - \Phi(-Z_n) + o_P(1)$ as obtained in the first part of this proof, the last expression becomes

$$\begin{aligned} & \left(\begin{array}{l} 1(\widehat{X}_{\text{med}} = 0)P^*(\widehat{X}_{\text{med}}^* \leq -1) + 1(\widehat{X}_{\text{med}} = 1)P^*(\widehat{X}_{\text{med}}^* \leq 0) \\ 1(\widehat{X}_{\text{med}} = 0)P^*(\widehat{X}_{\text{med}}^* \leq 0) + 1(\widehat{X}_{\text{med}} = 1)P^*(\widehat{X}_{\text{med}}^* \leq 1) - \frac{1}{2} \end{array} \right) + o_P(1) \\ &= \left(\begin{array}{l} 1(\frac{1}{2} < \Phi(-Z_n))P^*(\widehat{X}_{\text{med}}^* \leq 0) \\ 1(\frac{1}{2} \geq \Phi(-Z_n))P^*(\widehat{X}_{\text{med}}^* \leq 0) + 1(\frac{1}{2} < \Phi(-Z_n)) - \frac{1}{2} \end{array} \right) + o_P(1) \\ &= \left(\begin{array}{l} 1(\frac{1}{2} < \Phi(-Z_n))(1 - \Phi(-Z_n)) \\ 1(\frac{1}{2} \geq \Phi(-Z_n))(1 - \Phi(-Z_n)) + 1(\frac{1}{2} < \Phi(-Z_n)) - \frac{1}{2} \end{array} \right) + o_P(1), \end{aligned}$$

which converges in probability by the continuous mapping theorem [see e.g. Pollard (1984, III.6)] towards

$$\begin{aligned} & \left(\begin{array}{l} 1(\frac{1}{2} < \Phi(-Z))(1 - \Phi(-Z)) \\ 1(\frac{1}{2} \geq \Phi(-Z))(1 - \Phi(-Z)) + 1(\frac{1}{2} < \Phi(-Z)) - \frac{1}{2} \end{array} \right) \\ &= \left(\begin{array}{l} 1(\frac{1}{2} < \tilde{U})(1 - \tilde{U}) \\ 1(\frac{1}{2} \geq \tilde{U})(1 - \tilde{U}) + 1(\frac{1}{2} < \tilde{U}) - \frac{1}{2} \end{array} \right). \end{aligned} \tag{32}$$

Further, it holds

$$\begin{aligned} 1\left(\frac{1}{2} < \tilde{U}\right)(1 - \tilde{U}) &= \begin{cases} 1 - \tilde{U}, & \frac{1}{2} < \tilde{U} \leq \left\{\frac{1}{2}, \frac{1}{2} < \tilde{U}\right. \\ 0, & \frac{1}{2} \geq \tilde{U} \end{cases} \\ &= 1\left(\frac{1}{2} \geq \tilde{U}\right)(1 - \tilde{U}) + 1\left(\frac{1}{2} < \tilde{U}\right) - \frac{1}{2} \end{aligned}$$

such that the second component of (32) is always the maximum of both. To derive the cdf, let $x \in \mathbb{R}$ and, with $U = 1 - \tilde{U}$, we get

$$\begin{aligned}
 & P\left(1\left(\frac{1}{2} \geq \tilde{U}\right)(1 - \tilde{U}) + 1\left(\frac{1}{2} < \tilde{U}\right) - \frac{1}{2} \leq x\right) \\
 &= P\left(1\left(\frac{1}{2} \leq U\right)U + 1\left(\frac{1}{2} > U\right) - \frac{1}{2} \leq x\right) \\
 &= P\left(1\left(\frac{1}{2} \leq U\right)U + 1\left(\frac{1}{2} > U\right) - \frac{1}{2} \leq x, U \geq \frac{1}{2}\right) \\
 &\quad + P\left(1\left(\frac{1}{2} \leq U\right)U + 1\left(\frac{1}{2} > U\right) - \frac{1}{2} \leq x, U < \frac{1}{2}\right) \\
 &= P\left(\frac{1}{2} \leq U \leq x + \frac{1}{2}\right) + P\left(\frac{1}{2} \leq x, U < \frac{1}{2}\right) \\
 &= \begin{cases} 0, & x < 0 \\ x, & x \in [0, 1/2) \\ 1/2, & x \geq 1/2 \end{cases} + \begin{cases} 0, & x < 0 \\ 0, & x \in [0, 1/2) \\ 1/2, & x \geq 1/2 \end{cases} \\
 &= x1_{[0, \frac{1}{2})}(x) + 1_{[\frac{1}{2}, \infty)}(x).
 \end{aligned}$$

□

Proof of Theorem 2 (i) This part is a special case of Theorem 8.

(ii) The second statement follows similarly to the proof of Theorem 1 using part (i), the results from above and from

$$\begin{aligned}
 & \left(P^*(\hat{X}_{m,\text{med}}^* - \hat{X}_{\text{med}} \leq -1) - P(\hat{X}_{\text{med}} - X_{\text{med}} \leq -1) \right) \\
 & \left(P^*(\hat{X}_{m,\text{med}}^* - \hat{X}_{\text{med}} \leq 0) - P(\hat{X}_{\text{med}} - X_{\text{med}} \leq 0) \right) \\
 &= \left(1\left(\frac{1}{2} < \Phi(-Z_n)\right)P^*(\hat{X}_{m,\text{med}}^* \leq 0) \right. \\
 & \quad \left. + 1\left(\frac{1}{2} \geq \Phi(-Z_n)\right)P^*(\hat{X}_{m,\text{med}}^* \leq 0) + 1\left(\frac{1}{2} < \Phi(-Z_n)\right) - \frac{1}{2} \right) + o_P(1) \\
 &= \left(1\left(\frac{1}{2} < \tilde{U}\right)\frac{1}{2} \right. \\
 & \quad \left. + 1\left(\frac{1}{2} \geq \tilde{U}\right)\frac{1}{2} + 1\left(\frac{1}{2} < \tilde{U}\right) - \frac{1}{2} \right) + o_P(1) \\
 &= \left(1(U < \frac{1}{2})\frac{1}{2} \right. \\
 & \quad \left. + 1(U \geq \frac{1}{2})\frac{1}{2} + 1(U < \frac{1}{2}) - \frac{1}{2} \right) + o_P(1)
 \end{aligned}$$

as $1(U < 1/2)1/2 = 1(U \geq 1/2)1/2 + 1(U < 1/2) - 1/2$ and $1(U < 1/2) = 2\tilde{S}$ is Bernoulli-distributed. □

Proof of Theorem 6 (i) Note that \hat{Q}_p and Q_p take their values in V only. Under our assumptions on V there exists an $\epsilon > 0$ such that for each $p \in (0, 1)$

$$P(\hat{Q}_p = Q_p) = P(\hat{Q}_p \in (Q_p - \epsilon, Q_p + \epsilon]).$$

This implies

$$P(\widehat{Q}_p = Q_p) = P(p \leq \widehat{F}_n(Q_p + \epsilon)) - P(p \leq \widehat{F}_n(Q_p - \epsilon)) \tag{33}$$

due to the monotonicity of \widehat{F}_n . The first term on the rhs tends to one by the WLLN, which is a consequence of Theorem 20, and the second term vanishes asymptotically with the same reasoning.

- (ii) This follows from part (iii) with $d = k = 1$ and symmetry of a univariate, centered normal random variable.
- (iii) As we proved consistency of the sample quantiles if $F(Q_{p_i}) > p_i$ in (i), we can restrict the computations to the case where $F(Q_{p_i}) = p_i$ and $i = 1, \dots, k$ in the following. Similarly to (33), we get

$$\begin{aligned} P(\widehat{Q} = q) &= P\left(\underline{0} \in \times_{j=1}^k (\sqrt{n}(\widehat{F}_n(q_j - \epsilon) - F(Q_{p_j} + \epsilon)), \sqrt{n}(\widehat{F}_n(q_j + \epsilon) - F(Q_{p_j} + \epsilon)))\right) \\ &\xrightarrow{n \rightarrow \infty} P(\underline{0} \in \times_{j=1}^k (-\infty 1(q_j = Q_{p_j}) + Z_j 1(q_j = v_{l_{j+1}}), \infty 1(q_j = v_{l_{j+1}}) + Z_j 1(q_j = Q_{p_j}))) \\ &= P\left(\bigcap_{j=1}^k \{(2 \cdot 1(q_j = Q_{p_j}) - 1)Z_j \geq 0\}\right), \end{aligned} \tag{34}$$

where the multivariate CLT

$$\sqrt{n} \begin{pmatrix} \widehat{F}_n(Q_{p_1}) - F(Q_{p_1}) \\ \vdots \\ \widehat{F}_n(Q_{p_k}) - F(Q_{p_k}) \end{pmatrix} \xrightarrow{\mathcal{D}} \underline{Z} \sim \mathcal{N}(\underline{0}, \mathbf{W})$$

has been used; see Theorem 20. □

Proof of Theorem 8 It suffices to verify

$$\sup_{\underline{x} \in \mathbb{R}^d} \left| P^*(\widehat{Q}_m^* \leq \underline{x}) - P(\widehat{Q} \leq \underline{x}) \right| = \sup_{\underline{k} \in V^d} \left| P^*(\widehat{Q}_m^* \leq \underline{k}) - P(\widehat{Q} \leq \underline{k}) \right| = o_P(1)$$

due to the discrete nature of the underlying process. First we get from Theorem 6(i) above with $Q_{p_j} = v_{l_j}$ that

$$\lim_{n \rightarrow \infty} P(\widehat{Q}_{p_j} < Q_{p_j}) = \lim_{n \rightarrow \infty} P(\widehat{Q}_{p_j} > v_{l_{j+1}}) = 0.$$

Similarly, deducing a bootstrap WLLN from Lemma 21, we get

$$P^*(\widehat{Q}_{p_j,m}^* < Q_{p_j}) + P^*(\widehat{Q}_{p_j,m}^* > v_{l_{j+1}}) = o_P(1)$$

as well. Using the notation of Theorem 6(iii), it remains to show that

$$P^*(\widehat{Q}_m^* = \underline{q}) \xrightarrow{P} P\left(\bigcap_{j=1}^k \{(2 \cdot 1(q_j = Q_{p_j}) - 1)Z_j \geq 0\}\right).$$

Actually, we get

$$\begin{aligned} P^*(\widehat{Q}_m^* = \underline{q}) &= P^*\left(\underline{0} \in \times_{j=1}^k \left(\sqrt{m}(\widehat{F}_m^*(q_j - \epsilon) - F(Q_{p_j} + \epsilon)), \sqrt{m}(\widehat{F}_m^*(q_j + \epsilon) - F(Q_{p_j} + \epsilon))\right)\right) \\ &= P^*\left(\underline{0} \in \times_{j=1}^k \left(\sqrt{m}(\widehat{F}_m^*(q_j - \epsilon) - \widehat{F}_n(Q_{p_j} + \epsilon)) + O_P((m/n)^{1/2}), \right. \right. \\ &\quad \left. \left. \sqrt{m}(\widehat{F}_m^*(q_j + \epsilon) - \widehat{F}_n(Q_{p_j} + \epsilon)) + O_P((m/n)^{1/2})\right)\right). \end{aligned}$$

Further, as $m = o(n)$ and from Lemma 21 the last right-hand side converges in probability to $P\left(\bigcap_{j=1}^k \{(2 \cdot 1(q_j = Q_{p_j}) - 1)Z_j \geq 0\}\right)$, which proves bootstrap consistency. □

Proof of Theorem 9 The proof follows in analogy to the proof of Theorem 8 from Theorem 22. □

Proof of Theorem 11 For a specific $j \in \mathbb{Z}$ we have $Q_p = v_j$. From bootstrap consistency we obtain $P^*(\widehat{Q}_{p,m}^* \in [Q_p, v_{j+1}]) \xrightarrow{P} 1$ and $P^*(\widehat{Q}_{p,m}^* = v_{j+1}) \xrightarrow{P} 1/2$ or 0 if $F(Q_p) = p$ and $F(Q_p) > p$, respectively. Hence, $\text{cov}_L \xrightarrow{P} 1$.

Concerning the coverage of the small set we obtain

$$\text{cov}_S \xrightarrow{P} \begin{cases} 1/2 & \text{if } F(Q_p) = p \\ 0 & \text{if } F(Q_p) > p \end{cases}.$$

In particular, this implies that

$$P^* \xrightarrow{P} \begin{cases} 1 - 2\alpha & \text{if } F(Q_p) = p \\ 1 - \alpha & \text{if } F(Q_p) > p \end{cases}.$$

From Theorems 6, 8 and 9, we get $H(\widehat{Q}_{p,m}^*) = Q_p$ with probability tending to one. Noting that the difference between both coverages is larger than 1/4 with probability tending to one, we obtain

$$\begin{aligned} P(Q_p \in \text{CS}) &= P(\widehat{Q}_p \in \text{CS}_L \mathbf{1}(Y \leq p^*) + \text{CS}_S \mathbf{1}(Y > p^*), \text{cov}_L - \text{cov}_S \geq 1/4) \\ &\quad + o(1) \end{aligned}$$

$$\begin{aligned}
 &= E\left(p^*1(\widehat{Q}_p \in CS_L, \text{cov}_L - \text{cov}_S \geq 1/4) \right. \\
 &\quad \left. + (1 - p^*)1(\widehat{Q}_p \in CS_S, \text{cov}_L - \text{cov}_S \geq 1/4)\right) + o(1) \\
 &= (1 - \alpha)E\left(\frac{1}{\text{cov}_L - \text{cov}_S}1(\widehat{Q}_p \in CS_L \setminus CS_S, \text{cov}_L - \text{cov}_S \geq 1/4)\right) \\
 &\quad - E\left(\frac{\text{cov}_S}{\text{cov}_L - \text{cov}_S}1(\widehat{Q}_p \in CS_L, \text{cov}_L - \text{cov}_S \geq 1/4)\right) + o(1) \\
 &\quad + E\left(\frac{\text{cov}_L}{\text{cov}_L - \text{cov}_S}1(\widehat{Q}_p \in CS_S, \text{cov}_L - \text{cov}_S \geq 1/4)\right) + o(1) \\
 &=: P_1 + P_2 + P_3 + o(1).
 \end{aligned}$$

Moreover, it holds

$$P(\widehat{Q}_p \in CS_L) = P\left(\widehat{Q}_p \in \left[F_{\widehat{Q}_{p,m}^*}^{*-1}(\alpha/2), F_{\widehat{Q}_{p,m}^*}^{*-1}(1 - \alpha/2)\right]\right) \xrightarrow{n \rightarrow \infty} 1,$$

and

$$\begin{aligned}
 P(\widehat{Q}_p \in CS_S) &= P\left(\widehat{Q}_p \in \left[F_{\widehat{Q}_{p,m}^*}^{*-1}(\alpha/2), F_{\widehat{Q}_{p,m}^*}^{*-1}(1 - \alpha/2)\right]\right) \\
 &\xrightarrow{n \rightarrow \infty} \begin{cases} 1/2 & \text{if } F(Q_p) = p \\ 0 & \text{if } F(Q_p) > p \end{cases},
 \end{aligned}$$

and therefore

$$P(\widehat{Q}_p \in CS_L \setminus CS_S) \xrightarrow{n \rightarrow \infty} \begin{cases} 1/2 & \text{if } F(Q_p) = p \\ 1 & \text{if } F(Q_p) > p \end{cases}.$$

Bringing all together, we get from Theorem 25.11 in Billingsley (1995)

$$P_1 \xrightarrow{n \rightarrow \infty} 1 - \alpha, \quad P_2 \xrightarrow{n \rightarrow \infty} \begin{cases} -1 & \text{if } F(Q_p) = p \\ 0 & \text{if } F(Q_p) > p \end{cases}, \quad \text{and} \quad P_3 \xrightarrow{n \rightarrow \infty} \begin{cases} 1 & \text{if } F(Q_p) = p \\ 0 & \text{if } F(Q_p) > p \end{cases}$$

since the random variables whose expectations we calculate in P_1, \dots, P_3 are bounded by 4. This finally implies that CS has asymptotically exact level $1 - \alpha$. \square

Proof of Theorem 14 The proofs of the first two cases $p < F_{mid}(v_1)$ and $p > F_{mid}(v_d)$ follow the same lines. They can be carried out in complete analogy to the proofs of Theorem 2, Case 1 and Case 2 in Ma et al. (2011) if we can show that

$$\frac{1}{n} \sum_{t=1}^n 1(X_t = v_k) \xrightarrow{P} P(X_1 = v_k), \quad k = 1, \dots, d. \tag{35}$$

This in turn follows from the WLLN that can be deduced from Theorem 20 noting that $1(X_t = v_k) = 1(X_t \leq v_k) - 1(X_t \leq v_{k-1})$ for $k = 2, \dots, d$ and $1(X_t = v_1) = 1(X_t \leq v_1)$.

If $p = \lambda F_{\text{mid}}(v_{k+1}) + (1 - \lambda)F_{\text{mid}}(v_{k+2})$ such that $\lambda \in (0, 1)$, we can pursue the steps of the proof of Theorem 2, Case 3 in Ma et al. (2011) to get

$$\begin{aligned} & \sqrt{n} (\widehat{Q}_{p,\text{mid}} - Q_{p,\text{mid}}) \\ &= \sqrt{n} (v_{k+1} - v_{k+2}) \left[\frac{\widehat{F}_{\text{mid}}(v_{k+2}) - p}{\widehat{F}_{\text{mid}}(v_{k+2}) - \widehat{F}_{\text{mid}}(v_{k+1})} - \frac{F_{\text{mid}}(v_{k+2}) - p}{F_{\text{mid}}(v_{k+2}) - F_{\text{mid}}(v_{k+1})} \right]. \end{aligned} \tag{36}$$

Now, asymptotics of (36) can be deduced easily using the Δ -method, if we can show that

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n (Y_1 + \dots + Y_n) \xrightarrow{d} \mathcal{N}(0_d, \Sigma), \tag{37}$$

where $Y_t = (1(X_t = v_1) - P(X_t = v_1), \dots, 1(X_t = v_d) - P(X_t = v_d))'$, $t = 1, \dots, n$ and $\Sigma = (\Sigma_{j_1, j_2})_{j_1, \dots, j_2=1, \dots, d}$. Now using the same representation of the indicator functions as in the first part of the proof, (37) follows from Theorem 20 and the continuous mapping theorem. To this end, note that $\widehat{F}_{\text{mid}}(v_k) = n^{-1} \sum_{t=1}^n \{ \sum_{i=1}^k 1(X_t = v_i) - 0.5 1(X_t = v_k) \}$ and similarly $F_{\text{mid}}(v_k) = \sum_{i=1}^k a_i - 0.5 a_k$.

The assertion for the case $p = F_{\text{mid}}(v_{k+1})$, $k = 1, \dots, d - 2$ can be deduced from Theorem 20 in the same manner as in the proof of Theorem 2, Case 4 in Ma et al. (2011).

The proofs of the last two boundary cases $p = F_{\text{mid}}(v_1)$ and $p = F_{\text{mid}}(v_d)$ follow the same lines and we show only the first one. As $\sqrt{n}(\widehat{F}_{\text{mid}}(v_1) - F_{\text{mid}}(v_1)) = O_P(1)$ by Theorem 20, for sufficiently large n , there is a λ_n such that $0 < \lambda_n < 1$ and $p = \lambda_n \widehat{F}_{\text{mid}}(v_2) + (1 - \lambda_n)\widehat{F}_{\text{mid}}(v_1)$ if $\widehat{F}_{\text{mid}}(v_1) < p$. Then, from the definition of $\widehat{Q}_{p,\text{mid}}$, we get

$$\sqrt{n}(\widehat{Q}_{p,\text{mid}} - Q_{p,\text{mid}}) = \widetilde{Z}_n \frac{v_2 - v_1}{\widehat{F}_{\text{mid}}(v_2) - \widehat{F}_{\text{mid}}(v_1)} 1(0 < \widetilde{Z}_n), \tag{38}$$

where $\widetilde{Z}_n = \sqrt{n}(p - \widehat{F}_{\text{mid}}(v_1)) = \sqrt{n}(F_{\text{mid}}(v_1) - \widehat{F}_{\text{mid}}(v_1))$. From (38), we get

$$\begin{aligned} & P(\sqrt{n}(\widehat{Q}_{p,\text{mid}} - Q_{p,\text{mid}}) \leq x) \\ &= \begin{cases} 0, & x < 0 \\ P(\widetilde{Z}_n \geq 0), & x = 0 \\ P(\widetilde{Z}_n \geq 0) + P\left(\widetilde{Z}_n \frac{v_2 - v_1}{\widehat{F}_{\text{mid}}(v_2) - \widehat{F}_{\text{mid}}(v_1)} \in (0, x]\right), & x > 0 \end{cases} \end{aligned}$$

and the cases on the last right-hand side converge corresponding to the claimed limiting distribution again using Theorem 20. □

Proof of Theorem 17 To prove the first case of (29), let $\widehat{a}_j^* = m^{-1} \sum_{t=1}^m 1(X_t^* = j)$, $\widehat{a}_j = n^{-1} \sum_{t=1}^n 1(X_t = j)$, $j = 1, \dots, d$, and $\widehat{a}_0 = \widehat{a}_{d+1} = \widehat{a}_0^* = \widehat{a}_{d+1}^* = 0$. For sufficiently large n , with probability tending to 1 and because of $\sqrt{n}(\widehat{a}_j - a_j) = O_P(1)$ and $\sqrt{m}(\widehat{a}_j^* - \widehat{a}_j) = O_{P^*}(1)$ due to Lemma 21 and Theorem 22, we can find a λ_m^* with $0 < \lambda_m^* < 1$ such that $p = \lambda_m^* \widehat{a}_0^* + (1 - \lambda_m^*) \widehat{a}_1^*$. Consequently from (28), we get

$$\widehat{Q}_{p,\text{mid},m}^* = \lambda_m^* v_0 + (1 - \lambda_m^*) v_1 = v_1 = Q_{p,\text{mid}}$$

with probability tending to one as $v_0 = v_1$. By analogue arguments, we get also $\widehat{Q}_{p,\text{mid},m}^* = v_d = Q_{p,\text{mid}}$ with probability tending to one if $p > F_{\text{mid}}(v_d)$.

Similarly, for the second case of (29), with probability tending to one, we can find a λ_m^* with $0 < \lambda_m^* < 1$ such that $p = \lambda_m^* \widehat{F}_{\text{mid},m}^*(v_{k+1}) + (1 - \lambda_m^*) F_{\text{mid},m}^*(v_{k+2})$. Similar to (36), this leads to

$$\begin{aligned} & \sqrt{m}(\widehat{Q}_{p,\text{mid},m}^* - \widehat{Q}_{p,\text{mid}}) \\ &= (v_{k+1} - v_{k+2}) \left[\frac{\widehat{F}_{\text{mid},m}^*(v_{k+2}) - p}{\widehat{F}_{\text{mid},m}^*(v_{k+2}) - \widehat{F}_{\text{mid},m}^*(v_{k+1})} - \frac{\widehat{F}_{\text{mid},m}(v_{k+2}) - p}{\widehat{F}_{\text{mid},m}(v_{k+2}) - \widehat{F}_{\text{mid},m}(v_{k+1})} \right] \end{aligned}$$

which converges conditionally to the claimed normal distribution by Lemma 21 and Theorem 22 and by the Δ -method similar to the proof of Theorem 14.

To prove (30), as $\sqrt{m}(\widehat{F}_{\text{mid},m}^*(v_{k+1}) - \widehat{F}_{\text{mid}}(v_{k+1})) = O_{P^*}(1)$ by Lemma 21 and Theorem 22, we get similar to the proof of Case 4 of Theorem 2 in Ma et al. (2011) that

$$\begin{aligned} \widehat{F}_{\text{mid}}(v_{k+1}) &= \widehat{F}_{\text{mid},m}^*(v_{k+1}) + 1(\widehat{F}_{\text{mid},m}^*(v_{k+1}) \\ &\geq \widehat{F}_{\text{mid}}(v_{k+1})) \lambda_{m1}^* (\widehat{F}_{\text{mid},m}^*(v_k) - \widehat{F}_{\text{mid},m}^*(v_{k+1})) \\ &\quad + 1(\widehat{F}_{\text{mid},m}^*(v_{k+1}) < \widehat{F}_{\text{mid}}(v_{k+1})) \lambda_{m2}^* (\widehat{F}_{\text{mid},m}^*(v_{k+2}) - \widehat{F}_{\text{mid},m}^*(v_{k+1})) \end{aligned}$$

holds for some $0 \leq \lambda_{m1}^*, \lambda_{m2}^* < 1$. With $\widetilde{Z}_m^* = \sqrt{m}(\widehat{F}_{\text{mid}}(v_{k+1}) - \widehat{F}_{\text{mid},m}^*(v_{k+1}))$ and (28), this leads to

$$\begin{aligned} \widehat{Q}_{p,\text{mid},m}^* &= 1(0 \geq \widetilde{Z}_m^*) \{ \lambda_{m1}^* v_k + (1 - \lambda_{m1}^*) v_{k+1} \} \\ &\quad + 1(0 < \widetilde{Z}_m^*) \{ \lambda_{m2}^* v_{k+2} + (1 - \lambda_{m2}^*) v_{k+1} \} \\ &= v_{k+1} + \left(1(0 \geq \widetilde{Z}_m^*) \frac{v_{k+1} - v_k}{\widehat{F}_{\text{mid},m}^*(v_{k+1}) - \widehat{F}_{\text{mid},m}^*(v_k)} \right. \\ &\quad \left. + 1(0 < \widetilde{Z}_m^*) \frac{v_{k+2} - v_{k+1}}{\widehat{F}_{\text{mid},m}^*(v_{k+2}) - \widehat{F}_{\text{mid},m}^*(v_{k+1})} \right) \frac{\widetilde{Z}_m^*}{\sqrt{m}} \end{aligned}$$

and

$$\begin{aligned} \sqrt{m}(\widehat{Q}_{p,\text{mid},m}^* - Q_{p,\text{mid}}) &= \left(1(0 \geq \widetilde{Z}_m^*) \frac{v_{k+1} - v_k}{\widehat{F}_{\text{mid},m}^*(v_{k+1}) - \widehat{F}_{\text{mid},m}^*(v_k)} \right. \\ &\quad \left. + 1(0 < \widetilde{Z}_m^*) \frac{v_{k+2} - v_{k+1}}{\widehat{F}_{\text{mid},m}^*(v_{k+2}) - \widehat{F}_{\text{mid},m}^*(v_{k+1})} \right) \widetilde{Z}_m^*. \end{aligned}$$

Finally, we can show for all $x \in \mathbb{R}$ that

$$P^* \left(\sqrt{m}(\widehat{Q}_{p,\text{mid},m}^* - Q_{p,\text{mid}}) \leq x \right) \rightarrow \begin{cases} 0, & x < 0, k = 0 \\ P \left(\widetilde{Z} \leq x \frac{F_{\text{mid}}(v_{k+1}) - F_{\text{mid}}(v_k)}{v_{k+1} - v_k} \right), & x < 0, k > 0 \\ \frac{1}{2}, & x = 0, k \in \{0, 1, \dots, d - 2\} \\ \frac{1}{2} + P \left(\widetilde{Z} \in (0, x \frac{F_{\text{mid}}(v_{k+2}) - F_{\text{mid}}(v_{k+1})}{v_{k+2} - v_{k+1}}] \right), & x > 0, k < d - 1 \\ 1, & x \geq 0, k = d - 1 \end{cases}$$

in probability, where $\widetilde{Z} \sim \mathcal{N}(0, \sigma_{\widetilde{Z}}^2)$ and

$$\begin{aligned} \sigma_{\widetilde{Z}}^2 &= \sum_{h \in \mathbb{Z}} \text{cov} \left(\sum_{j=1}^{k+1} 1(X_h \leq v_j) - 0.5 \cdot 1(X_h = v_{k+1}), \right. \\ &\quad \left. \sum_{j=1}^{k+1} 1(X_0 \leq v_{k+1}) - 0.5 \cdot 1(X_0 = v_{k+1}) \right) \end{aligned}$$

can be obtained from Lemma 21 and Theorem 22, respectively. This concludes this proof. □

Proof of Corollary 18 First it follows from Theorem 17 and (31) that the distribution of $\sqrt{m}(\widehat{Q}_{p,\text{mid},m}^* - \widehat{Q}_{p,\text{mid}})$ converges in probability to the same limit as the distribution of $\sqrt{n}(\widehat{Q}_{p,\text{mid}}^* - Q_{p,\text{mid}})$, i.e. either to zero or to one of the distributions of Z_1 to Z_4 . To prove convergence of the corresponding distribution functions in the Kolmogorov-Smirnov metric we treat the different cases separately. First, let $p < F_{\text{mid}}(v_1)$ (or $p > F_{\text{mid}}(v_d)$) which can be considered in the same manner and hence, the proof is omitted). From Remark 15(i) we obtain

$$\begin{aligned} &\sup_{x \in \mathbb{R}} \left| P^*(\sqrt{m}(\widehat{Q}_{p,\text{mid},m}^* - \widehat{Q}_{p,\text{mid}}) \leq x) - P(\sqrt{n}(\widehat{Q}_{p,\text{mid}} - Q_{p,\text{mid}}) \leq x) \right| \\ &\leq \sup_{x < 0} \left| P^*(\sqrt{m}(\widehat{Q}_{p,\text{mid},m}^* - \widehat{Q}_{p,\text{mid}}) \leq x) - P(\sqrt{n}(\widehat{Q}_{p,\text{mid}} - Q_{p,\text{mid}}) \leq x) \right| \\ &\quad + 1 - P^*(\sqrt{m}(\widehat{Q}_{p,\text{mid},m}^* - \widehat{Q}_{p,\text{mid}}) \leq 0) + 1 - P(\sqrt{n}(\widehat{Q}_{p,\text{mid}} - Q_{p,\text{mid}}) \leq 0) \\ &\leq \lim_{x \uparrow 0} P^*(\sqrt{m}(\widehat{Q}_{p,\text{mid},m}^* - \widehat{Q}_{p,\text{mid}}) \leq x) + \lim_{x \uparrow 0} P(\sqrt{n}(\widehat{Q}_{p,\text{mid}} - Q_{p,\text{mid}}) \leq x) + o_P(1) \end{aligned}$$

$$\begin{aligned} &\leq P^*(\widehat{Q}_{p,\text{mid},m}^* < \widehat{Q}_{p,\text{mid}}) + P(\widehat{Q}_{p,\text{mid}} < Q_{p,\text{mid}}) + o_P(1) \\ &= o_P(1). \end{aligned}$$

In the second and third case, i.e. when the limiting distribution is Z_1 or Z_2 , Polya’s theorem can be applied to deduce convergence in the Kolmogorov–Smirnov metric from distributional convergence since the limiting distribution function is continuous. It remains to consider the $p = F_{\text{mid}}(v_1)$ and $p = F_{\text{mid}}(v_d)$. Since they are similar again, we focus on the first setup. With the same arguments as in the proof of Polya’s theorem we get

$$\begin{aligned} &\sup_{x \in \mathbb{R}} \left| P^*(\sqrt{m}(\widehat{Q}_{p,\text{mid},m}^* - \widehat{Q}_{p,\text{mid}}) \leq x) - P(\sqrt{n}(\widehat{Q}_{p,\text{mid}} - Q_{p,\text{mid}}) \leq x) \right| \\ &\leq \sup_{x \leq 0} \left| P^*(\sqrt{m}(\widehat{Q}_{p,\text{mid},m}^* - \widehat{Q}_{p,\text{mid}}) \leq x) - P(\sqrt{n}(\widehat{Q}_{p,\text{mid}} - Q_{p,\text{mid}}) \leq x) \right| \\ &\quad + o_P(1). \end{aligned}$$

Now, we proceed similarly to the first case and, finally, we get

$$\begin{aligned} &\sup_{x \leq 0} \left| P^*(\sqrt{m}(\widehat{Q}_{p,\text{mid},m}^* - \widehat{Q}_{p,\text{mid}}) \leq x) - P(\sqrt{n}(\widehat{Q}_{p,\text{mid}} - Q_{p,\text{mid}}) \leq x) \right| \\ &\leq \left| P^*(\widehat{Q}_{p,\text{mid},m}^* < \widehat{Q}_{p,\text{mid}}) - P(\widehat{Q}_{p,\text{mid}} < Q_{p,\text{mid}}) \right| \\ &\quad + \frac{1}{2} - P^*(\sqrt{m}(\widehat{Q}_{p,\text{mid},m}^* - \widehat{Q}_{p,\text{mid}}) \leq 0) + \frac{1}{2} - P(\sqrt{n}(\widehat{Q}_{p,\text{mid}} - Q_{p,\text{mid}}) \leq 0) \\ &= o_P(1). \end{aligned}$$

□

Proof of Theorem 19 First, note that

$$\begin{aligned} P(Q_{p,\text{mid}} \in \text{CI}) &= P\left(\sqrt{n}(\widehat{Q}_{p,\text{mid}} - Q_{p,\text{mid}}) \in \text{CI}_{S,\text{mid}}^{(r)} \cap \text{CI}_{S,\text{mid}}^{(l)} \right. \\ &\quad + \text{CI}_{L,\text{mid}} \setminus (\text{CI}_{S,\text{mid}}^{(r)} \cap \text{CI}_{S,\text{mid}}^{(l)}) \mathbf{1}(Y \leq p_{\text{mid}}^*) \\ &\quad + \text{CI}_{L,\text{mid}} \setminus \text{CI}_{S,\text{mid}}^{(l)} \mathbf{1}(Y > p_{\text{mid}}^*, \text{cov}_{S,\text{mid}}^{(r)} \leq 1 - \alpha) \\ &\quad \left. + \text{CI}_{L,\text{mid}} \setminus \text{CI}_{S,\text{mid}}^{(r)} \mathbf{1}(Y > p_{\text{mid}}^*, \text{cov}_{S,\text{mid}}^{(r)} > 1 - \alpha)\right), \end{aligned}$$

where $+$ above indicates the disjoint union. We consider the rhs in a case-by-case manner.

The cases $p < F_{\text{mid}}(v_1)$ and $p > F_{\text{mid}}(v_d)$ can be treated similarly and we only give the calculations for the first setup. Here, $\text{cov}_{L,\text{mid}} \xrightarrow{P} 1$ and $\text{cov}_{S,\text{mid}}^{(r)} \xrightarrow{P} 0$ which then implies that $p_{\text{mid}}^* \xrightarrow{P} 1 - \alpha$. Now the proof can be carried out in complete analogy to the proof of Theorem 11 (case of $F(Q_p) > p$).

Also the cases where Z_1 and Z_2 are the limiting variables have a similar structure which results from continuity of the corresponding limiting cdf's. Here, we get

$$P(Q_{p,\text{mid}} \in \text{CI}) = P\left(\sqrt{n}(\widehat{Q}_{p,\text{mid}} - Q_{p,\text{mid}}) \in \text{CI}_{S,\text{mid}}^{(r)} \cap \text{CI}_{S,\text{mid}}^{(l)}\right) + o_P(1)$$

where the latter probability tends to $1 - \alpha$ as $n \rightarrow \infty$.

Next we consider the case $p = F_{\text{mid}}(v_1)$. Since α is assumed to be less than $1/2$ and the limiting distribution has a normal density on the positive half line, $\text{cov}_{S,\text{mid}}^{(r)} \xrightarrow{\mathcal{P}} 1 - \alpha/2$ which in turn implies

$$\begin{aligned} P(Q_{p,\text{mid}} \in \text{CI}) &= P\left(\sqrt{n}(\widehat{Q}_{p,\text{mid}} - Q_{p,\text{mid}}) \in \text{CI}_{S,\text{mid}}^{(r)} \cap \text{CI}_{S,\text{mid}}^{(l)}\right) \\ &\quad + \text{CI}_{L,\text{mid}} \setminus (\text{CI}_{S,\text{mid}}^{(r)} \cap \text{CI}_{S,\text{mid}}^{(l)}) \mathbf{1}(Y \leq p_{\text{mid}}^*) \\ &\quad + \text{CI}_{L,\text{mid}} \setminus \text{CI}_{S,\text{mid}}^{(r)} \mathbf{1}(Y > p_{\text{mid}}^*, \text{cov}_{S,\text{mid}}^{(r)} > 1 - \alpha) + o(1). \end{aligned}$$

Since $p_{\text{mid}}^* \xrightarrow{\mathcal{P}} 1 - \alpha$, it can be shown in analogy to the proof of Theorem 11 that $P(Q_{p,\text{mid}} \in \text{CI}) \xrightarrow[n \rightarrow \infty]{} 1 - \alpha$. It remains to investigate the case $p = F_{\text{mid}}(v_d)$.

Here, $\text{cov}_{L,\text{mid}} \xrightarrow{\mathcal{P}} 1 - \alpha/2$ and $\text{cov}_{S,\text{mid}}^{(r)} \xrightarrow{\mathcal{P}} 1/2 - \alpha/2$ which then implies that $p_{\text{mid}}^* \xrightarrow{\mathcal{P}} 1 - \alpha$. The desired result follows with the same arguments as before. \square

Auxiliary results

Theorem 20 (CLT under τ -dependence) *Suppose that $(X_t)_{t \in \mathbb{Z}}$ is a τ -dependent process with $\sum_{h=0}^{\infty} \tau(h) < \infty$. Then for all $x_1, \dots, x_D \in \mathbb{R}$, $D \in \mathbb{N}$,*

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n (1(X_t \leq x_1) - F(x_1), \dots, 1(X_t \leq x_D) - F(x_D))' \xrightarrow{\mathcal{D}} \mathcal{N}(\underline{0}, \mathbf{W})$$

with

$$\mathbf{W} = \left(\sum_{h \in \mathbb{Z}} \text{cov}(1(X_h \leq x_{j_1}), 1(X_0 \leq x_{j_2})) \right)_{j_1, j_2=1, \dots, D}.$$

Proof We apply the multivariate central limit theorem for weakly dependent data of Leucht and Neumann (2013, Theorem 6.1). To this end, we check its prerequisites with $Z_t := (1(X_t \leq x_1) - F(x_1), \dots, 1(X_t \leq x_D) - F(x_D))' / \sqrt{n}$. Obviously, these variables are centered and $\sum_{t=1}^n E \|Z_t\|_2^2 < \infty$. Also the Lindeberg condition clearly holds true by stationarity and boundedness of the underlying process $(X_t)_{t \in \mathbb{Z}}$. Next we have to show that

$$\left[\text{cov} \left(\sum_{t=1}^n Z_t \right) \right]_{j_1, j_2} \xrightarrow{n \rightarrow \infty} W_{j_1, j_2}.$$

We consider the component-wise absolute difference between both terms

$$\begin{aligned} & \left| \frac{1}{n} \sum_{s,t=1}^n \text{cov}(1(X_s \leq x_{j_1}), 1(X_t \leq x_{j_2})) - \sum_{h \in \mathbb{Z}} \text{cov}(1(X_h \leq x_{j_1}), 1(X_0 \leq x_{j_2})) \right| \\ & \leq \sum_{h \in \mathbb{Z}} \min \left\{ \frac{|h|}{n}, 1 \right\} |\text{cov}(1(X_h \leq x_{j_1}), 1(X_0 \leq x_{j_2}))| \end{aligned}$$

which converges to zero by dominated convergence theorem if $\sum_{h \in \mathbb{Z}} |\text{cov}(1(X_h \leq x_{j_1}), 1(X_0 \leq x_{j_2}))| < \infty$. This in turn can be deduced from the presumed summability of the τ -coefficients if $|\text{cov}(1(X_h \leq x_{j_1}), 1(X_0 \leq x_{j_2}))| \leq \text{const. } \tau(h)$. To see this, first note that for any $\nu < \min_k \{v_{k+1} - v_k\}$, and for $v_k \leq x < v_{k+1}$

$$1(X_1 \leq x) = 1(X_1 \leq v_k) = 1(X_1 \leq v_k + \nu) - \frac{X_1 - v_k}{\nu} 1(v_k \leq X_1 \leq v_k + \nu) \quad \text{a.s.}$$

where the rhs is a Lipschitz continuous function in X_1 . Now we use coupling arguments to obtain an upper bound for the absolute values of the covariances under consideration when $h > 0$. The case $h < 0$ can be treated similarly and is therefore omitted. Let \tilde{X}_h denote a copy of X_h that is independent of X_0 and such that $E|\tilde{X}_h - X_h| \leq \tau(h)$. With $x_{j_1} \in [v_k, v_{k+1})$ for a suitable k , we obtain

$$\begin{aligned} & |\text{cov}(1(X_h \leq x_{j_1}), 1(X_0 \leq x_{j_2}))| \\ & \leq E \left| 1(X_h \leq x_{j_1}) - 1(\tilde{X}_h \leq x_{j_1}) \right| \\ & \leq E \left| 1(X_h \leq v_k + \nu) - \frac{X_h - v_k}{\nu} 1(v_k \leq X_h \leq v_k + \nu) - 1(\tilde{X}_h \leq v_k + \nu) \right. \\ & \quad \left. + \frac{\tilde{X}_h - v_k}{\nu} 1(v_k \leq \tilde{X}_h \leq v_k + \nu) \right| \\ & \leq \frac{1}{\nu} E|\tilde{X}_h - X_h| \\ & \leq \frac{\tau(h)}{\nu}. \end{aligned} \tag{39}$$

Finally we have to check two conditions of weak dependence. Let $g: \mathbb{R}^{du} \rightarrow \mathbb{R}$ be a measurable function with $\|g\|_\infty \leq 1$ and $1 \leq s_1 < s_2 < \dots < s_u < s_u + h = t_1 \leq t_2 \in \mathbb{N}$. Again, in analogy to (39), we obtain

$$\text{cov}(g(Z_{s_1}, \dots, Z_{s_u})Z_{s_u, j_1}, Z_{t_1, j_2}) \leq \frac{1}{\nu n} \tau(t_1 - s_u),$$

which implies condition (6.27) with $\theta_h = \tau(h)/\nu$ in [Leucht and Neumann \(2013\)](#). Validity of their condition (6.28) follows from

$$\text{cov}(g(Z_{s_1}, \dots, Z_{s_u}), Z_{t_1, j_1} Z_{t_2, j_2}) \leq \frac{4}{\nu n} \tau(t_1 - s_u),$$

which completes the proof of the multivariate CLT. □

Lemma 21 (Bootstrap analogue to Theorem 20 for i.i.d. data) *Suppose that $(X_t)_{t \in \mathbb{Z}}$ is a sequence of i.i.d. random variables. Let X_1^*, \dots, X_m^* be drawn independently from \widehat{F}_n . Suppose that $m \rightarrow \infty$ and $m = o(n)$ or $m = n$. Then, for all $x_1, \dots, x_D \in \mathbb{R}$, $D \in \mathbb{N}$,*

$$\frac{1}{\sqrt{m}} \sum_{t=1}^m (1(X_t^* \leq x_1) - \widehat{F}_n(x_1), \dots, 1(X_t^* \leq x_D) - \widehat{F}_n(x_D))' \xrightarrow{D} \mathcal{N}(\underline{0}, \mathbf{W})$$

in probability, where

$$\mathbf{W} = (\text{cov}(1(X_0 \leq x_{j_1}), 1(X_0 \leq x_{j_2})))_{j_1, j_2=1, \dots, D}.$$

Proof This is an immediate consequence of Theorem 2.2 in [Bickel and Friedman \(1981\)](#). □

Theorem 22 (Block bootstrap analogue to Theorem 20) *Suppose that the assumptions of Theorem 20 hold true and that $\sum_{h=1}^{\infty} h \tau(h) < \infty$. Let X_1^*, \dots, X_m^* be an m -out-of- n block bootstrap sample. Suppose that $l/m + 1/l = o(1)$ as well as $m = o(n)$ or $m = n$ as $n \rightarrow \infty$. Then, for all $x_1, \dots, x_D \in \mathbb{R}$, $D \in \mathbb{N}$,*

$$\frac{1}{\sqrt{m}} \sum_{k=1}^m (1(X_k^* \leq x_1) - \widehat{F}_n(x_1), \dots, 1(X_k^* \leq x_D) - \widehat{F}_n(x_D))' \xrightarrow{D} \mathcal{N}(\underline{0}, \mathbf{W})$$

in probability, where

$$\mathbf{W} = \left(\sum_{h \in \mathbb{Z}} \text{cov} (1(X_h \leq x_{j_1}), 1(X_0 \leq x_{j_2})) \right)_{j_1, j_2=1, \dots, D}.$$

Proof For notational convenience, we suppose $m = lb$ and let us introduce the notation

$$Z_k^* = \frac{1}{\sqrt{m}} (1(X_k^* \leq x_1) - \widehat{F}_n(x_1), \dots, 1(X_k^* \leq x_D) - \widehat{F}_n(x_D))',$$

$$\widetilde{Z}_k^* = \frac{1}{\sqrt{m}} (1(X_k^* \leq x_1) - E^*(1(X_k^* \leq x_1)), \dots, 1(X_k^* \leq x_D) - E^*(1(X_k^* \leq x_D)))',$$

such that it suffices to show $\sum_{k=1}^m (Z_k^* - \tilde{Z}_k^*) = o_{P^*}(1)$ and $\sum_{k=1}^m \tilde{Z}_k^* \xrightarrow{D} \mathcal{N}(\mathbf{0}, \mathbf{W})$ in probability. Considering the first part component-wise, for all j , we get

$$\begin{aligned} \sum_{k=1}^m (Z_k^* - \tilde{Z}_k^*)_j &= \frac{1}{\sqrt{m}} \sum_{k=1}^m (E^*(1(X_k^* \leq x_j)) - \hat{F}_n(x_j)) \\ &= \sqrt{m} \left(\frac{1}{n-l+1} \sum_{t=1}^n 1(X_t \leq x_j) - \frac{1}{n} \sum_{t=1}^n 1(X_t \leq x_j) \right) \\ &\quad + \frac{\sqrt{m}}{n-l+1} \sum_{t=1}^{l-1} \frac{t-l}{l} 1(X_t \leq x_j) \\ &\quad + \frac{\sqrt{m}}{n-l+1} \sum_{t=n-l+2}^n \frac{n-l+1-t}{l} 1(X_t \leq x_j) \\ &= A_1 + A_2 + A_3. \end{aligned}$$

Taking unconditional expectation of the last right-hand side gives a zero such that it suffices to show $A_i - E(A_i) = o_P(1)$ for $i = 1, 2, 3$. For the first term, we get from Theorem 20 that

$$\begin{aligned} A_1 - E(A_1) &= \frac{\sqrt{m}}{\sqrt{n}} \frac{l-1}{n-l+1} \left(\frac{1}{\sqrt{n}} \sum_{t=1}^n (1(X_t \leq x_j) - E(1(X_t \leq x_j))) \right) \\ &= O_P \left(\frac{\sqrt{m} l}{\sqrt{n} n} \right) \end{aligned}$$

vanishes as $l = o(m)$ by assumption. For the second term, we obtain

$$\begin{aligned} \text{var}(A_2) &= \frac{m}{(n-l+1)^2} \sum_{t_1, t_2=1}^{l-1} \frac{(t_1-l)(t_2-l)}{l^2} \text{cov}(1(X_{t_1} \leq x_j), 1(X_{t_2} \leq x_j)) \\ &\leq \frac{ml}{(n-l+1)^2} \sum_{h=-(l-2)}^{l-2} \frac{1}{l} \sum_{t=\max(1, 1-h)}^{\min(l-1, l-1-h)} \left| \frac{(h+t-l)(t-l)}{l^2} \right| \\ &\quad \cdot |\text{cov}(1(X_{h+t} \leq x_j), 1(X_t \leq x_j))| \\ &= O \left(\frac{ml}{n^2} \right) \end{aligned}$$

since the covariances are summable by $\sum_{h=1}^\infty \tau(h) < \infty$; see also (39) for details. Hence A_2 vanishes under the same conditions for l and m as for term A_1 above. The arguments for A_3 are completely analogue and we omit the details. To prove the (conditional) CLT along the lines of Section 4.2.2 in [Wieczorek \(2014\)](#) for

$$\sum_{k=1}^m \tilde{Z}_k^* = \sum_{r=1}^b \left(\sum_{s=(r-1)l+1}^{rl} \tilde{Z}_s^* \right) =: \sum_{r=1}^b \tilde{Y}_r^*,$$

observe that $\{\tilde{Y}_r^*, r = 1, \dots, b\}$ forms a triangular array of (conditionally) i.i.d. random variables with $E^*(\tilde{Y}_r^*) = 0$ by construction. Further, for [the (j_1, j_2) -component of] the conditional covariance, we have

$$\begin{aligned} \left[\text{cov}^* \left(\sum_{r=1}^b \tilde{Y}_r^* \right) \right]_{j_1, j_2} &= \sum_{r=1}^b \sum_{s_1, s_2=(r-1)l+1}^{rl} \text{cov}^* \left(\tilde{Z}_{s_1, j_1}^*, \tilde{Z}_{s_2, j_2}^* \right) \\ &= \frac{1}{l} \sum_{s_1, s_2=1}^l \text{cov}^* (1(X_{s_1}^* \leq x_{j_1}), 1(X_{s_2}^* \leq x_{j_2})) \\ &= \frac{1}{l} \sum_{s_1, s_2=1}^l \left(\frac{1}{n-l+1} \sum_{t=0}^{n-l} 1(X_{t+s_1} \leq x_{j_1}) 1(X_{t+s_2} \leq x_{j_2}) \right) \\ &\quad - \frac{1}{l} \left(\frac{1}{n-l+1} \sum_{t_1=0}^{n-l} \sum_{s_1=1}^l 1(X_{t_1+s_1} \leq x_{j_1}) \right) \\ &\quad \times \left(\frac{1}{n-l+1} \sum_{t_2=0}^{n-l} \sum_{s_2=1}^l 1(X_{t_2+s_2} \leq x_{j_2}) \right). \\ &= \frac{1}{l} \sum_{s_1, s_2=1}^l \left(\frac{1}{n-l+1} \sum_{t=0}^{n-l} T_{t+s_1, j_1} T_{t+s_2, j_2} \right) \\ &\quad - \left(\frac{1}{\sqrt{l}(n-l+1)} \sum_{t_1=0}^{n-l} \sum_{s_1=1}^l T_{t_1+s_1, j_1} \right) \\ &\quad \times \left(\frac{1}{\sqrt{l}(n-l+1)} \sum_{t_2=0}^{n-l} \sum_{s_2=1}^l T_{t_2+s_2, j_2} \right) \\ &=: I_1 - I_2 * I_3, \end{aligned}$$

where we have set $T_{t, j} = 1(X_t \leq x_j) - P(X_t \leq x_j)$. The terms I_2 and I_3 behave similarly and we only consider I_2 . Since $E I_2 = 0$, we show $I_2 = o_P(1)$ by proving that its variance vanishes asymptotically. We get

$$\begin{aligned} \text{var}(I_2) &\leq \frac{1}{n-l+1} \sum_{h_1=-(n-l)}^{n-l} \sum_{h_2=-(l-1)}^{l-1} \left(\frac{n-l+1-|h_1|}{n-l+1} \right) \left(\frac{l-|h_2|}{l} \right) \\ &\quad \times |\text{cov}(1(X_{h_1+h_2} \leq j_1), 1(X_0 \leq j_1))| \end{aligned}$$

which is of order $O(l/n)$ by (39).

By taking unconditional expectation of the first term I_1 , we obtain

$$E(I_1) = \sum_{h=-(l-1)}^{l-1} \frac{l-|h|}{l} \text{cov}(1(X_h \leq x_{j_1}), 1(X_0 \leq x_{j_2}))$$

and, by dominated convergence, the latter tends to W_{j_1, j_2} as desired. Hence, it remains to show that $\text{var}(I_1) = o(1)$ holds. By rewriting the arising covariances in terms of cumulants, we get

$$\begin{aligned} \text{var}(I_1) &= \frac{1}{l^2(n-l+1)^2} \sum_{s_1, s_2, s_3, s_4=1}^l \sum_{t_1, t_2=0}^{n-l} \text{cov}(T_{t_1+s_1, j_1} T_{t_1+s_2, j_2}, T_{t_2+s_3, j_1} T_{t_2+s_4, j_2}) \\ &= \frac{1}{l^2(n-l+1)^2} \sum_{s_1, s_2, s_3, s_4=1}^l \sum_{t_1, t_2=0}^{n-l} \left\{ E(T_{t_1+s_1, j_1} T_{t_2+s_3, j_1}) E(T_{t_1+s_2, j_2} T_{t_2+s_4, j_2}) \right. \\ &\quad \left. + E(T_{t_1+s_1, j_1} T_{t_2+s_4, j_2}) E(T_{t_1+s_2, j_2} T_{t_2+s_3, j_1}) \right. \\ &\quad \left. + \text{cum}(T_{t_1+s_1, j_1}, T_{t_1+s_2, j_2}, T_{t_2+s_3, j_1}, T_{t_2+s_4, j_2}) \right\}, \end{aligned}$$

where we have used that $\text{cum}(A, B, C, D) = E(ABCD) - E(AB)E(CD) - E(AC)E(BD) - E(AD)E(BC)$ for centered random variables A, B, C, D holds. As $E(T_{t_1+s_1, j_1} T_{t_2+s_3, j_1}) = \text{cov}(1(X_{t_1+s_1} \leq x_{j_1}), 1(X_{t_2+s_3} \leq x_{j_1})) \leq C\tau(|t_1+s_1-t_2-s_3|)$, by invoking the covariance inequality (39), the first and second summands on the rhs above can shown to be of order $O(l/n)$.

Next, we establish an upper bound bound for $|\text{cum}(T_{t_1, j_1}, T_{t_2, j_2}, T_{t_3, j_3}, T_{t_4, j_4})|$, where we assume w.l.o.g. that $t_1 \leq \dots \leq t_4$. Let $R = \max\{t_4 - t_3, t_3 - t_2, t_2 - t_1\}$. We consider each of the three possible values of R separately. First, suppose that $R = t_4 - t_3$. Then using the same coupling techniques as in the proof of Theorem 20, we get similarly to (39)

$$\begin{aligned} |\text{cum}(T_{t_1, j_1}, T_{t_2, j_2}, T_{t_3, j_3}, T_{t_4, j_4})| &\leq C\tau(R) [1 + |ET_{t_1, j_1} T_{t_2, j_2}| + |ET_{t_1, j_1} T_{t_3, j_3}| \\ &\quad + |ET_{t_2, j_2} T_{t_3, j_3}|] \\ &\leq 4C\tau(R) \end{aligned} \tag{40}$$

with some finite constant C since $\|T_{t_i, j_i}\|_\infty \leq 1$. If $R = t_3 - t_2$, we obtain

$$\begin{aligned} |\text{cum}(T_{t_1, j_1}, T_{t_2, j_2}, T_{t_3, j_3}, T_{t_4, j_4})| &\leq \text{cov}(T_{t_1, j_1} T_{t_2, j_2}, T_{t_3, j_3} T_{t_4, j_4}) \\ &\quad + C\tau(R) [|ET_{t_2, j_2} T_{t_4, j_4}| + |ET_{t_1, j_1} T_{t_4, j_4}|] \\ &\leq 4C\tau(R). \end{aligned} \tag{41}$$

Finally, in case of $R = t_2 - t_1$ the cumulant can be bounded as follows

$$\begin{aligned}
 |\text{cum}(T_{t_1, j_1}, T_{t_2, j_2}, T_{t_3, j_3}, T_{t_4, j_4})| &\leq 3C\tau(R) + C\tau(R) [|ET_{t_3, j_3} T_{t_4, j_4}| \\
 &\quad + |ET_{t_2, j_2} T_{t_4, j_4}| + |ET_{t_2, j_2} T_{t_3, j_3}|] \\
 &\leq 6C\tau(R).
 \end{aligned}
 \tag{42}$$

To sum up, we obtain

$$\begin{aligned}
 \text{var}(I_1) &\leq \frac{1}{l^2(n-l+1)^2} \sum_{s_1, s_2, s_3, s_4=1}^l \sum_{t_1, t_2=0}^{n-l} |\text{cum}(T_{t_1+s_1, j_1}, T_{t_1+s_2, j_2}, T_{t_2+s_3, j_1}, T_{t_2+s_4, j_2})| + o(1) \\
 &\leq \frac{6Cl}{n} \sum_{h=1}^{n-l} h\tau(h) + o(1),
 \end{aligned}$$

which vanishes asymptotically since we assumed $\sum_{h=1}^{\infty} h\tau(h) < \infty$.

To complete the proof of the bootstrap CLT, it remains to show the Lindeberg condition to be able to apply (a multivariate version of) Lindeberg-Feller’s CLT for independent triangular arrays. That is, as $\text{cov}^*(\sum_{r=1}^b \tilde{Y}_r^*) = O_P(1)$ holds by the calculations above, for all $\epsilon > 0$, it remains to show

$$\sum_{r=1}^b E^* \left(\|\tilde{Y}_r^*\|_2^2 \mathbf{1}(\|\tilde{Y}_r^*\|_2 \geq \epsilon) \right) = b E^* \left(\|\tilde{Y}_1^*\|_2^2 \mathbf{1}(\|\tilde{Y}_1^*\|_2 \geq \epsilon) \right) = o_P(1)$$

as $\{\tilde{Y}_r^*, r = 1, \dots, b\}$ forms a triangular array of (conditionally) i.i.d. random variables. Computing the conditional expectation leads to

$$b E^* \left(\|\tilde{Y}_1^*\|_2^2 \mathbf{1}(\|\tilde{Y}_1^*\|_2 \geq \epsilon) \right) = \frac{b}{n-l+1} \sum_{t=0}^{n-l} \left\| \sum_{s=1}^l \tilde{Z}_{s+t} \right\|_2^2 \mathbf{1} \left(\left\| \sum_{s=1}^l \tilde{Z}_{s+t} \right\|_2 \geq \epsilon \right),
 \tag{43}$$

where

$$\begin{aligned}
 \tilde{Z}_{t+s} &= \frac{1}{\sqrt{m}} (1(X_{t+s} \leq x_1) - E^*(1(X_s^* \leq x_1)), \dots, 1(X_{t+s} \leq x_D) \\
 &\quad - E^*(1(X_s^* \leq x_D)))'
 \end{aligned}$$

with $E^*(1(X_s^* \leq x_1)) = \frac{1}{n-l+1} \sum_{t_1=0}^{n-l} 1(X_{t_1+s} \leq x_1)$. Now, we want to replace $\|\sum_{s=1}^l \tilde{Z}_{s+t}\|_2^2$ by $\|\sum_{s=1}^l Z_{s+t}\|_2^2$, where

$$Z_{t+s} = \frac{1}{\sqrt{m}} (1(X_{t+s} \leq x_1) - F(x_1), \dots, 1(X_{t+s} \leq x_D) - F(x_D))',$$

which leads to the upper bound

$$\begin{aligned} & \frac{2b}{n-l+1} \sum_{t=0}^{n-l} \left\| \sum_{s=1}^l Z_{s+t} \right\|_2^2 \mathbb{1} \left(\left\| \sum_{s=1}^l \tilde{Z}_{s+t} \right\|_2 \geq \epsilon \right) \\ & + \frac{2b}{n-l+1} \sum_{t=0}^{n-l} \left\| \sum_{s=1}^l (\tilde{Z}_{s+t} - Z_{s+t}) \right\|_2^2 \mathbb{1} \left(\left\| \sum_{s=1}^l \tilde{Z}_{s+t} \right\|_2 \geq \epsilon \right) \\ & =: II_1 + II_2 \end{aligned}$$

for (43). Considering the second summand above component-wise, it is straightforward to show that for all j , it holds

$$\sum_{s=1}^l (Z_{s+t} - \tilde{Z}_{s+t}) = \frac{1}{\sqrt{m}(n-l+1)} \sum_{s=1}^l \sum_{t_1=0}^{n-l} (T_{t_1+s,1}, \dots, T_{t_1+s,D})'$$

which is independent of t , such that with $T_t = (T_{t,1}, \dots, T_{t,D})'$

$$II_2 \leq \frac{2b}{m} \left\| \frac{1}{n-l+1} \sum_{s=1}^l \sum_{t_1=0}^{n-l} T_{t_1+s} \right\|_2^2 = O_P \left(\frac{l}{n} \right)$$

by the same arguments as used before to treat I_2 . Concerning II_1 , as all summands are non-negative, it suffices to show $E|II_1| = E(II_1) = o(1)$. From stationarity and by application of Cauchy–Schwarz inequality, we get

$$\begin{aligned} E(II_1^2) &= E \left(2b \left\| \sum_{s=1}^l Z_s \right\|_2^2 \mathbb{1} \left(\left\| \sum_{s=1}^l \tilde{Z}_s \right\|_2 \geq \epsilon \right) \right)^2 \\ &\leq 4b^2 E \left(\left\| \sum_{s=1}^l Z_s \right\|_2^4 \right) P \left(\left\| \sum_{s=1}^l \tilde{Z}_s \right\|_2 \geq \epsilon \right). \end{aligned}$$

As the second factor above is vanishing by Markov inequality and since $E(\|\sum_{s=1}^l \tilde{Z}_s\|_2^2) = O(l/m)$, it remains to show that $b^2 E(\|\sum_{s=1}^l Z_{s,j}\|_4^4) = O(1)$ for all j . Rewriting things in terms of cumulants, we get

$$\begin{aligned} b^2 E \left(\left\| \sum_{s=1}^l \tilde{Z}_{s,j} \right\|_4^4 \right) &= b^2 \sum_{s_1, s_2, s_3, s_4=1}^l E(Z_{s_1,j} Z_{s_2,j} Z_{s_3,j} Z_{s_4,j}) \\ &= 3 \left(\frac{1}{l} \sum_{s_1, s_2=1}^l \text{cov}(1(X_{s_1} \leq x_j), 1(X_{s_2} \leq x_j)) \right)^2 \\ &\quad + \frac{1}{l^2} \sum_{s_1, s_2, s_3, s_4=1}^l \text{cum}(T_{s_1,j}, T_{s_2,j}, T_{s_3,j}, T_{s_4,j}) \end{aligned}$$

as higher-order cumulants are invariant to shifts. The first summand on the last rhs is uniformly bounded. The second summand is also of order $O(1)$ by (40) to (42) and $\sum_{h=1}^{\infty} h \tau(h) < \infty$. \square

Acknowledgments This research was supported by the Research Center (SFB) 884 “Political Economy of Reforms” (Project B6), funded by the German Research Foundation (DFG). The authors are grateful to Tobias Niebuhr, Technische Universität Braunschweig, for fruitful discussions that motivated this project and his assistance with the implementation of the numerical examples. The authors thank two anonymous referees for their careful reading and insightful suggestions.

References

- Angus, J. E. (1993). Asymptotic theory for bootstrapping the extremes. *Communications in Statistics—Theory and Methods*, 22, 15–30.
- Athreya, K. B., Fukuchi, J. (1994). Bootstrapping extremes of i.i.d random variables. In J. Galambos, J. Lechner, E. Simiu (Eds.), *Proceedings of the Conference on Extreme Value Theory and Applications*, NIST Special Publication 866 (vol. 3). New York: Springer.
- Athreya, K. B., Fukuchi, J. (1997). Confidence intervals for endpoints of a c.d.f. via bootstrap. *Journal of Statistical Planning and Inference*, 58, 299–320.
- Athreya, K. B., Fukuchi, J., Lahiri, S. N. (1999). On the bootstrap and the moving block bootstrap for the maximum of a stationary process. *Journal of Statistical Planning and Inference*, 76(1–2), 1–17.
- Bickel, P. J., Friedman, D. A. (1981). Some asymptotic theory for the bootstrap. *The Annals of Statistics*, 9, 1196–1217.
- Bickel, P. J., Sakov, A. (2008). On the choice of m in the m out of n bootstrap and confidence bounds for extrema. *Statistica Sinica*, 18, 967–985.
- Billingsley, P. (1995). *Probability and measure*. New York: Wiley.
- Chen, J., Lazar, N. A. (2010). Quantile estimation for discrete data via empirical likelihood. *Journal of Nonparametric Statistics*, 22, 237–255.
- Dedecker, J., Prieur, C. (2004). Couplage pour la distance minimale. *Comptes Rendus Mathématique*, 338, 805–808.
- Dedecker, J., Prieur, C. (2005). New dependence coefficients. Examples and applications to statistics. *Probability Theory and Related Fields*, 132, 203–236.
- Deheuvels, P., Mason, D., Shorack, G. (1993). Some results on the influence of extremes on the bootstrap. *Annales de l’Institut Henri Poincaré*, 29, 83–103.
- Del Barrio, E., Janssen, A., Pauly, M. (2013). The $m(n)$ out of $k(n)$ bootstrap for partial sums of St. Petersburg type games. *Electronic Communications in Probability*, 18, 1–10.
- Doukhan, P., Fokianos, K., Li, X. (2012a). On weak dependence conditions: the case of discrete valued processes. *Statistics and Probability Letters*, 82, 1941–1948.
- Doukhan, P., Fokianos, K., Tjøstheim, D. (2012b). On weak dependence conditions for Poisson autoregressions. *Statistics and Probability Letters*, 82, 942–948.
- Drost, F. C., van den Akker, R., Werker, B. J. M. (2009). Efficient estimation of auto-regression parameters and innovation distributions for semiparametric integer-valued AR(p) models. *Journal of the Royal Statistical Society, Series B*, 71, 467–485.
- Efron, B. (1979). Bootstrap: another look at the jackknife. *The Annals of Statistics*, 7, 1–26.
- Ferland, R., Latour, A., Oraichi, D. (2006). Integer count GARCH processes. *Journal of Time Series Analysis*, 27, 923–942.
- Fokianos, K. (2011). Some recent progress in count time series. *Statistics: A Journal of Theoretical and Applied Statistics*, 45, 49–58.
- Fokianos, K., Rahbek, A., Tjøstheim, D. (2009). Poisson autoregression. *Journal of the American Statistical Association—Theory and Methods*, 104, 1430–1439.
- Harrell, F. E., Davis, C. E. (1982). A new distribution-free quantile estimator. *Biometrika*, 62, 635–640.
- Horowitz, J. (2001). The bootstrap. In: J. J. Heckman, E. E. Leamer (Ed.), *Handbook of Econometrics 5, chapter 52* (pp. 3159–3228). Elsevier.
- Krantz, S. G. (1991). *Real analysis and foundations*. Boca Raton: CRC Press.

- Leucht, A., Neumann, M. H. (2013). Dependent wild bootstrap for degenerate U- and V-statistics. *Journal of Multivariate Analysis*, *117*, 257–280.
- Ma, Y., Genton, M. G., Parzen, E. (2011). Asymptotic properties of sample quantiles of discrete distributions. *Annals of the Institute of Statistical Mathematics*, *63*, 227–243.
- Mammen, E. (1992). *When does the bootstrap work?: Asymptotic results and simulations*. New York, Heidelberg: Springer.
- McKenzie, E. (1988). ARMA models for dependent sequences of Poisson counts. *Advances in Applied Probability*, *20*, 822–835.
- Parzen, E. (1997). Concrete statistics. In S. Ghosh, W. R. Schucany, W. B. Smith (Eds.), *Statistics in quality* (pp. 309–332). New York: Marcel Dekker.
- Parzen, E. (2004). Quantile probability and statistical data modeling. *Statistical Science*, *19*, 652–662.
- Pollard, D. (1984). *Convergence of stochastic processes*. New York: Springer.
- Resnick, S. I. (1987). *Extreme values, regular variation, and point processes*. New York: Springer.
- Santana, L. (2009). Contributions to the m -out-of- n bootstrap. Dissertation. North-West University, Potchefstroom Campus, South Africa.
- Serfling, R. J. (2002). *Approximation theorems of mathematical statistics*. New York: Wiley.
- Shao, J., Chen, Y. (1998). Bootstrapping sample quantiles based on complex survey data under hot deck imputation. *Statistica Sinica*, *8*, 1071–1086.
- Sharipov, O. S. H., Wendler, M. (2013). Normal limits, nonnormal limits, and the bootstrap for quantiles of dependent data. *Statistics and Probability Letters*, *83*, 1028–1035.
- Sun, S., Lahiri, S. N. (2006). Bootstrapping the Sample Quantile of a Weakly Dependent Sequence. *Shankya*, *68*, 130–166.
- Swanepoel, J. W. H. (1986). A note on proving that the (modified) bootstrap works. *Communications in Statistics - Theory and Methods*, *15*, 3193–3203.
- Tempelmeier, H. (2000). Inventory service-levels in customer supply chain. *OR Spektrum*, *22*, 361–380.
- Thas, O., De Neve, J., Clement, L., Otooy, J.-P. (2012). Probabilistic index models. *Journal of the Royal Statistical Society, Series B74, Part 4*, 623–671.
- Wang, D., & Hutson, A. D. (2011). A fractional order statistic towards defining a smooth quantile function for discrete data. *Journal of Statistical Planning and Inference*, *141*, 3142–3150.
- Weiß, C. H. (2008). Thinning operations for modeling time series of counts—a survey. *Advances in Statistical Analysis*, *92*, 319–341.
- Wieczorek, B. (2014). Blockwise bootstrap of the estimated empirical process based on ψ -weakly dependent observations (Submitted).