



Article Bootstrapping Not Independent and Not Identically Distributed Data

Martin Hrba¹, Matúš Maciak¹, Barbora Peštová² and Michal Pešta^{1,*}

- ¹ Department of Probability and Mathematical Statistics, Faculty of Mathematics and Physics, Charles University, Sokolovská 49/83, 18675 Prague, Czech Republic
- ² Department of Statistical Modelling, Institute of Computer Science, Czech Academy of Sciences, Pod Vodárenskou věží 271/2, 18207 Prague, Czech Republic
- Correspondence: michal.pesta@mff.cuni.cz

Abstract: Classical normal asymptotics could bring serious pitfalls in statistical inference, because some parameters appearing in the limit distributions are unknown and, moreover, complicated to estimated (from a theoretical as well as computational point of view). Due to this, plenty of stochastic approaches for constructing confidence intervals and testing hypotheses cannot be directly applied. Bootstrap seems to be a plausible alternative. A methodological framework for bootstrapping not independent and not identically distributed data is presented together with theoretical justification of the proposed procedures. Among others, bootstrap laws of large numbers and central limit theorems are provided. The developed methods are utilized in insurance and psychometry.

Keywords: bootstrap; statistical inference; asymptotic normality; weakly dependent data; not identically distributed data; moving block bootstrap; law of large numbers; central limit theorem; psychometric evaluation; non-life insurance

MSC: 62E20; 62F40; 62P05; 62P15; 91G05

JEL Classification: C15; C18; C22; C53; G22

1. Introduction and Motivation

Stochastic inference frequently relies on the limiting normal distribution, which allows the characterization of large sample (asymptotic) behavior of the considered statistics. However, some entities of the asymptotic distribution such as the variance-covariance matrix quite often depend on the parameters or other quantities, which cannot be estimated directly from the data. In addition, this estimation can be performed neither theoretically nor computationally. Bootstrap methods—a *class of resampling techniques*—serve as suitable alternatives. Additional problems arise, when the underlying data do not form a random sample, simply because they cannot be considered as independent or identically distributed. Thus, the bootstrap techniques developed for independent and identically distributed (IID) observations need to be adjusted and suited up for the not independent and not identically distributed (NINID) data. The main goal of this paper is to postulate a theoretical setup and a methodological framework, which enable us to show how the bootstrap procedures should be accommodated for the NINID observations and, simultaneously, to formally justify the validity of such approaches. It is prerequisite to realize that there are two sources of randomness when bootstrapping: the first randomness comes from the data themselves, the second randomness is brought by the chance from random resampling. Therefore, we need to access, quantify, and control both sources.

However, it is not sufficient to give just the algorithmic procedures for bootstrapping the parameters of interest. It is also desired to possess fundamental mathematical guaranties that the particular bootstrap technique provides reasonable results. A possible way how



Citation: Hrba, M.; Maciak, M.; Peštová, B.; Pešta, M. Bootstrapping Not Independent and Not Identically Distributed Data. *Mathematics* **2022**, *10*, 4671. https://doi.org/10.3390/ math10244671

Academic Editor: Christophe Chesneau

Received: 15 November 2022 Accepted: 6 December 2022 Published: 9 December 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). to postulate such validity is to formulate proper *stochastic convergences*, assuring that the bootstrapped parameter is distributionally close to the estimator of the unknown parameter of interest.

1.1. State of the Art

The bootstrap was introduced by Efron in [1]. This theoretical concept was consequently investigated by [2] in the case of independent and identically distributed data. Monographs [3] or [4] provide a bootstrap "first" course overview. A comprehensive theoretical summary for the bootstrap methods can be found in [5]. Various types of the dependent bootstrap procedures were proposed and [6] provided a extensive summary. Subsampling within the bootstrap framework was given by [7]. The block bootstrap approach was independently suggested by [8,9] for the case of sample mean. [10,11] extended these results, but still considered only strictly stationary processes. Reference [12] generalized previous approaches for non-stationary processes. Bootstrapping panel data with non-stationary, heteroscedastic, and outlying observations was elaborated by [13]. Permutation bootstrap for the exchangeable observations was summarized by [14]. An application of the dependent block bootstrap approach for the regression setup was performed in [15,16]. In contrast to this theoretical asymptotic result, the practical finite sample properties for the dependent bootstrap were studied by [17–19].

1.2. Structure of the Paper

The organization of the paper is as follows. The bootstrap procedures for IID and NINID data are summarized in Section 2. Consequently, Section 3 postulates asymptotic relations and closeness in the bootstrap world. Various bootstrap laws of large numbers are derived in Section 4. The validity of the considered bootstrap methods are guaranteed by the bootstrap central limit theorems provided in Section 5. Practical applications to the problems from insurance and psychometry are shown in Section 6. Conclusions are relegated to Section 7, while all the proofs are collected in the Appendix A.

2. Bootstrap Methods for NINID Data

A brief summary for bootstrapping the IID data is going to be provided. After that, an extension for the NINID observations is given.

2.1. Independent Bootstrap

An *independent bootstrap* requires no distributional assumptions and, moreover, independence of the input multivariate observations $\{\xi_i\}_{i=1}^n$ is assumed while resampling is being performed. It is also often called the *nonparametric bootstrap* and it refers to the simplest scheme of independent resampling from the original observations $[\xi_1, \dots, \xi_n]$. The main idea behind the independent bootstrap lies in the resampling of the independent column data ξ_i 's with replacement in order to obtain the *bootstrapped data* ξ_i^* . Now, an entity of interest is calculated from the new "starred" data ξ_i^* 's, e.g., a test statistic or an estimator of an unknown parameter. It is desired that the new distribution of the bootstrapped entity *mimics* the original distribution of the concerned statistics. From now on, we concentrate on the bootstrap for the mean parameter. However, the presented approaches can be generalized for other parameters of interest as well, but this goes beyond the scope of this paper.

Assuming identical distribution and finite variance of ξ_i 's, the central limit theorem holds for the sample mean

$$\bar{\boldsymbol{\xi}}_n := \frac{1}{n} \sum_{i=1}^n \boldsymbol{\xi}_i$$

$$\bar{\boldsymbol{\xi}}_n^* := \frac{1}{n} \sum_{i=1}^n \boldsymbol{\xi}_i^*.$$

Afterwards, it is necessary to asymptotically compare the distributions of

$$\sqrt{n}(\bar{\boldsymbol{\xi}}_n - \mathbb{E}\boldsymbol{\xi}_1)$$
 and $\sqrt{n}(\bar{\boldsymbol{\xi}}_n^* - \bar{\boldsymbol{\xi}}_n)$

In a proper mathematical way to be sure that the empirical distribution of the bootstrap estimate $\bar{\xi}_n^*$ of the mean can be used instead of the distribution of $\bar{\xi}_n$. The asymptotic closeness of $\sqrt{n}(\bar{\xi}_n - \mathbb{E}\xi_1)$ and $\sqrt{n}(\bar{\xi}_n^* - \bar{\xi}_n)$ will be clarified and proved later on in Section 5. An algorithm for the independent bootstrap is shown in Procedure 1 and its validity will be proved in Theorem 6.

Procedure 1 Independent bootstrap for the sample mean.

Input: Data consisting of *n* IID vectors of observations ξ_i .

Output: Empirical bootstrap distribution of $\bar{\xi}_n$, i.e., the empirical distribution, where the probability mass 1/B concentrates at each of ${}_{(1)}\bar{\xi}_n^*, \dots, {}_{(B)}\bar{\xi}_n^*$.

- 1: calculate the sample mean $\bar{\xi}_n$
- 2: for b = 1 to B do // repeat in order to obtain empirical distribution of $\bar{\xi}_n$
- 3: $[(b)\xi_1^*, \dots, (b)\xi_n^*]$ resampled with replacement from the columns of $[\xi_1, \dots, \xi_n]$
- 4: re-calculate $_{(b)}\bar{\xi}_n^* \leftarrow n^{-1}\sum_{i=1}^n {}_{(b)}\xi_i^*$

It is silently supposed that the bootstrapped sample is of the same size as the original one, i.e., $\{\xi_i\}_{i=1}^n$ and $\{\xi_i^*\}_{i=1}^n$. Generally, one may consider resampled bootstrapped data $\{\xi_i^*\}_{i=1}^m$ consisting of *m* resamples from the data $\{\xi_i\}_{i=1}^n$ having sample size *n*. Thereby, an additional condition needs to be postulated on the rate of the sample sizes:

$$m = \mathcal{O}(n), \quad n \to \infty \qquad \& \qquad n = \mathcal{O}(m), \quad m \to \infty.$$

There is, however, no need to distinguish between the same sample size and two different sample sizes (the original one and the bootstrap one), which are asymptotically equivalent, at least from the theoretical asymptotic point of view. However, there could be some computational improvements when considering various samples size of the resampled data, which is not the target of this research.

2.2. Moving Block Bootstrap

Intuitively, a problem will arise when considering dependent and/or not identically distributed observations. A generalization of the independent bootstrap (case sampling with replacement)—*block bootstrap*—can be used. Instead of sampling individual cases, the *blocks of adjacent observations* are resampled with replacement. Stacking individual adjacent cases together into one solid block partly preserves the dependence between consecutive observations. Since the "weak" dependence can be seen as an asymptotic independence, the blocks can be resampled independently. It is a way how to achieve that the dependence between faraway observations is vanishing.

A plethora of types of block bootstrap techniques were suggested. A comprehensive summary is provided by [6]. In general, the way of drawing the blocks of observations defines the difference among the block bootstrap versions. Bootstrapping blocks, which do not overlap, is performed in the *non-overlapping block bootstrap*. Since some observations are not allowed to be joined into the same block, this approach is less efficient for estimation [20]. Therefore, the *moving block bootstrap* (MBB) is considered here. The key idea of the MBB is forming the consecutive block of observations from the previous one through shifting the

"stacking window" by one observation ahead. The MBB technique for the sample mean of multivariate observations is described in Procedure 2 in detail.

Procedure 2 Moving block bootstrap for the sample mean.

Input: Data consisting of *n* NINID vectors of observations ξ_i and n = mb.

Output: Empirical bootstrap distribution of sample mean $\bar{\xi}_n$, i.e., the empirical distribution, where the probability mass 1/D concentrates at each of $(1)\bar{\xi}_n^*, \dots, (D)\bar{\xi}_n^*$.

- 1: define \mathbf{B}_j as the block of *b* consecutive $\boldsymbol{\xi}_i$'s starting from $\boldsymbol{\xi}_j$, that is $\mathbf{B}_j = [\boldsymbol{\xi}_j, \dots, \boldsymbol{\xi}_{j+b-1}]$ for $j = 1, \dots, q$, where q := n b + 1
- 2: for d = 1 to D do // repeat in order to obtain empirical distribution of $\bar{\xi}_n$
- 3: resample with replacement ${}_{(d)}\mathbf{C}_1, \ldots, {}_{(d)}\mathbf{C}_m$ independently from $\{\mathbf{B}_1, \ldots, \mathbf{B}_q\}$ with equal probability 1/q, where each ${}_{(d)}\mathbf{C}_i$, $i = 1, \ldots, m$, is a block of size *b* with ${}_{(d)}\mathbf{C}_i = [{}_{(d)}\mathbf{c}_{i1}, \ldots, {}_{(d)}\mathbf{c}_{ib}]$
- 4: the MBB resample of size *n*, denoted by ${}_{(d)}\xi_1^*, \dots, {}_{(d)}\xi_n^*$, is formed by joining the ${}_{(d)}\mathbf{C}_1, \dots, {}_{(d)}\mathbf{C}_m$ to one big block, i.e., ${}_{(d)}\xi_i^* = {}_{(d)}\mathbf{c}_{\tau\nu}$ for $\tau = [(i-1)/b] + 1$, $\nu = i - b\tau$, and $i = 1, \dots, n / / {}_{(d)}\mathbf{\Xi}^* \equiv [{}_{(d)}\xi_1^*, \dots, {}_{(d)}\xi_n^*]$ is called the MBB version of $\mathbf{\Xi} \equiv [\xi_1, \dots, \xi_n]^\top$.
- 5: let the resample average be $_{(d)}\bar{\xi}_n^* \leftarrow n^{-1}\sum_{i=1}^n {}_{(d)}\xi_i^*$

```
6: end for
```

Let \mathbb{P}^* be the (bootstrap) distribution of ${}_{(d)}C_i$ conditional on the sample $\{\xi_1, \ldots, \xi_n\}$. So, given ξ_1, \ldots, ξ_n , the *m* random blocks, ${}_{(d)}C_1, \ldots, {}_{(d)}C_m$, are IID distributed according to \mathbb{P}^* . The length of the blocks—*blocksize*—is denoted by $b \in \mathbb{N}$. Without loss of generality to the asymptotic properties, let us suppose that $b \mid n$, i.e., there exist $m \in \mathbb{N}$ such that n = mb. In other words, we just neglect an integer division problem. For practical and computational purposes, if $b \nmid n$, then we can truncate the quotient n/b to an integer value, see [21].

An extension of the MBB is a *circular block bootstrap*, where the observations are not ordered on a single line, but they are put into a circle. The order of the observations is preserved with the only exception that the last observation on the circle is followed by the first one. Hence, the stacking window can join the first and the last observations into one block. The application of the circular block bootstrap as an extension of the MBB is postponed for some further work and is not considered in this paper.

2.3. Blocksize

A length of the blocks–the blocksize *b*–in the MBB procedure is a crucial decision. It indeed affects the bootstrapped statistics and the consequent statistical inference. The blocksize can be therefore considered as a *nuisance*, but also a *tuning parameter*. It will be derived in the forthcoming theory that $b = o(n^{1/2})$ as $n \to \infty$. However, for the practical and computational purposes, this asymptotic choice of *b* is rather cumbersome. With respect to the simulation studies from [12], it may be concluded that the blocksize choice $b = O(n^{1/3})$ as $n \to \infty$ could be asymptotically optimal according to the minimal mean square error of the moving block bootstrap variance estimator.

3. Types of Bootstrap Convergences

We would like to show that $\sqrt{n}(\bar{\xi}_n - \mathbb{E}\xi_1)$ and $\sqrt{n}(\bar{\xi}_n^* - \bar{\xi}_n)$ asymptotically coincide. This means that using the bootstrap distribution is no worse than using the asymptotic normal approximation. However, it does not mean that the bootstrap distribution better approximates the finite sample distribution of $\sqrt{n}(\bar{\xi}_n - \mathbb{E}\xi_1)$. To state this mathematically, we firstly need some formalization. For the rest of this paper, we link $\|\cdot\|$ together with any unitarily invariant matrix norm, e.g., the Frobenius matrix norm (and, in case of vectors, with the Euclidean norm).

Suppose that $\{\xi_n, \xi_n^*, \zeta_n, \zeta_n^*, \chi_n\}_{n=1}^{\infty}$ are sequences of random vectors/matrices, which elements exist on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The components of these sequences do not necessarily have to have the same dimension, e.g., ζ_n and ζ_{n+1} can have different dimensions for some $n \in \mathbb{N}$. Let us define a *conditional probability* given ζ_n

$$\mathbb{P}^*_{\boldsymbol{\zeta}_n}[\cdot] := \mathbb{E}_{\mathbb{P}}[\mathcal{I}(\cdot)|\boldsymbol{\zeta}_n]$$

Definition 1. Conditional weak convergence almost surely and in probability to each other. Let $\{\xi_n, \xi_n^*, \zeta_n\}_{n=1}^{\infty}$ be sequences of random vectors/matrices. If for every real-valued bounded continuous function *f* holds

$$\mathbb{E}[f(\boldsymbol{\xi}_n^*)] - \mathbb{E}[f(\boldsymbol{\xi}_n)] \xrightarrow[n \to \infty]{} 0,$$

then ξ_n^* and ξ_n are said to be approaching each other in distribution. In short, we write

$$\boldsymbol{\xi}_n^* \xleftarrow{\mathscr{D}}_{n \to \infty} \boldsymbol{\xi}_n.$$

if for every real-valued bounded continuous function f holds

$$\mathbb{E}[f(\boldsymbol{\xi}_n^*)|\boldsymbol{\zeta}_n] - \mathbb{E}[f(\boldsymbol{\xi}_n)] \xrightarrow[n \to \infty]{[\mathbb{P}]^{-a.s.}} 0,$$

then ξ_n^* conditioned on ζ_n and ξ_n are said to be approaching each other in distribution $[\mathbb{P}]$ -almost surely along ζ_n . In short, we write

$$\boldsymbol{\xi}_n^* | \boldsymbol{\zeta}_n \xleftarrow{\mathscr{D}([\mathbb{P}]\text{-}a.s.)}{n \to \infty} \boldsymbol{\xi}_n.$$

if for every real-valued bounded continuous function f holds

$$\mathbb{E}[f(\boldsymbol{\xi}_n^*)|\boldsymbol{\zeta}_n] - \mathbb{E}[f(\boldsymbol{\xi}_n)] \xrightarrow[n \to \infty]{\mathbb{P}} 0,$$

then ξ_n^* conditioned on ζ_n and ξ_n are said to be approaching each other in distribution in probability \mathbb{P} along ζ_n . In short, we write

$$\boldsymbol{\xi}_n^* \big| \boldsymbol{\zeta}_n \xleftarrow{\mathscr{D}(\mathbb{P})}{n \to \infty} \boldsymbol{\xi}_n.$$

In the same manner as above, we may define the distributional convergence on the "conditional" (resampled) level to a random variable ξ_0 ("constant" law).

Definition 2. Conditional weak convergence almost surely and in probability to a constant law. Let $\{\boldsymbol{\xi}_n^*, \boldsymbol{\zeta}_n\}_{n=1}^{\infty}$ be sequences of random vectors/matrices and $\boldsymbol{\xi}_0$ be a random vector/matrix. If for every real-valued bounded continuous function f holds

$$\mathbb{E}[f(\boldsymbol{\xi}_n^*)|\boldsymbol{\zeta}_n] \xrightarrow[n \to \infty]{[\mathbb{P}]^{\text{-}a.s.}} \mathbb{E}[f(\boldsymbol{\xi}_0)],$$

then ξ_n^* conditioned on ζ_n is said to converge to ξ_0 in distribution $[\mathbb{P}]$ -almost surely along ζ_n . In short, we write

$$\boldsymbol{\xi}_n^* \big| \boldsymbol{\zeta}_n \xrightarrow{\mathscr{D}([\mathbb{P}] \text{-} a.s.)}{n \to \infty} \boldsymbol{\xi}_0$$

if for every real-valued bounded continuous function f holds

$$\mathbb{E}[f(\boldsymbol{\xi}_n^*)|\boldsymbol{\zeta}_n] \xrightarrow[n \to \infty]{\mathbb{P}} \mathbb{E}[f(\boldsymbol{\xi}_0)],$$

then ξ_n^* conditioned on ζ_n is said to converge to ξ_0 in distribution in probability \mathbb{P} along ζ_n . In short, we write

$$\boldsymbol{\xi}_n^* \big| \boldsymbol{\zeta}_n \xrightarrow{\mathscr{D}(\mathbb{P})}{n \to \infty} \boldsymbol{\xi}_0.$$

Approaching in distribution to each other is often called *weakly approaching* to each other (almost surely or in probability along some sequence). The portmanteau lemma [22] ensures that Definition 1 is indeed appropriate. Hence, it gives equivalent characterizations for the weak convergence in distribution. To define convergence in probability \mathbb{P}^* , it means, convergences on the conditional level, two types of convergence in probability \mathbb{P}^* are defined: convergence in probability \mathbb{P} and convergence in [\mathbb{P}]-almost surely.

Definition 3. Convergence in conditional probability. Let $\{\xi_n, \xi_n^*, \zeta_n\}_{n=1}^{\infty}$ be sequences of random vectors/matrices. To say that $\xi_n^* - \xi_n$ converges in probability $\mathbb{P}_{\xi_n}^*$ to zero $[\mathbb{P}]$ -almost surely as n tends to infinity, i.e., $\xi_n^* - \xi_n \xrightarrow{\mathbb{P}_{\xi_n}^*([\mathbb{P}]^{-a.s.})} \mathbf{0},$

means

$$\forall \epsilon > 0: \mathbb{P}\Big[\lim_{n \to \infty} \mathbb{P}^*_{\zeta_n}[\|\boldsymbol{\xi}_n^* - \boldsymbol{\xi}_n\| \ge \epsilon] = 0\Big] = 1.$$
(1)

to say that $\xi_n^* - \xi_n$ converges in probability $\mathbb{P}_{\zeta_n}^*$ to zero in probability \mathbb{P} as n tends to infinity, i.e.,

$$\boldsymbol{\xi}_n^* - \boldsymbol{\xi}_n \xrightarrow{\mathbb{P}_{\boldsymbol{\zeta}_n}^*(\mathbb{P})} \mathbf{0}$$

means

$$\forall \epsilon > 0, \, \forall \tau > 0: \lim_{n \to \infty} \mathbb{P} \Big[\mathbb{P}^*_{\boldsymbol{\zeta}_n} [\| \boldsymbol{\xi}_n^* - \boldsymbol{\xi}_n \| \ge \epsilon] \ge \tau \Big] = 0.$$
⁽²⁾

Alternatively, (1) and (2) from Definition 3 can be read as

$$\forall \epsilon > 0: \left\{ \mathbb{P}^*_{\zeta_n}[\|\boldsymbol{\xi}_n^* - \boldsymbol{\xi}_n\| \ge \epsilon] \xrightarrow[n \to \infty]{[\mathbb{P}]^{-a.s.}} 0 \right\}$$

and

$$orall \epsilon > 0: \left\{ \mathbb{P}^*_{\zeta_n} [\| oldsymbol{\xi}^*_n - oldsymbol{\xi}_n \| \geq \epsilon] \xrightarrow[n o \infty]{\mathbb{P}} 0
ight\},$$

respectively.

In the same manner as above, we may define the convergence in probability \mathbb{P}^* on the resampled level to a random variable ξ_0 (does not depend on *n*).

Definition 4. Convergence in conditional probability to a variable. Let $\{\xi_n^*, \zeta_n\}_{n=1}^{\infty}$ be sequences of random vectors/matrices and ξ_0 is a random vector/matrix. To say that ξ_n^* converges to ξ_0 in probability $\mathbb{P}_{\xi_n}^*$ [P]-almost surely as n tends to infinity, i.e.,

$$\boldsymbol{\xi}_n^* \xrightarrow[n \to \infty]{\mathbb{P}_{\boldsymbol{\zeta}_n}^*([\mathbb{P}]\text{-}a.s.)} \boldsymbol{\xi}_0,$$

means that $\xi_n^* - \xi_0$ converges in probability $\mathbb{P}_{\xi_n}^*$ to zero $[\mathbb{P}]$ -almost surely as n tends to infinity. To say that ξ_n^* converges to ξ_0 in probability $\mathbb{P}_{\xi_n}^*$ in probability \mathbb{P} as n tends to infinity, i.e.,

$$\boldsymbol{\xi}_n^* \xrightarrow[n \to \infty]{\mathbb{P}_{\boldsymbol{\zeta}_n}^*(\mathbb{P})} \boldsymbol{\xi}_0,$$

means that $\xi_n^* - \xi_0$ converges in probability $\mathbb{P}_{\zeta_n}^*$ to zero in probability \mathbb{P} as n tends to infinity.

3.1. Properties of the Bootstrap Convergences

Important results concerning previously defined types of convergences summarized in the forthcoming will play a crucial role later on. Prokhorov's theorem can be extended into our setup.

Lemma 1. Assume that $\{\xi_n\}_{n=1}^{\infty}$ is tight. Then the following statements are equivalent: (i)

$$\boldsymbol{\xi}_n^* \big| \boldsymbol{\zeta}_n \xleftarrow{\mathscr{D}(\mathbb{P})}{n \to \infty} \boldsymbol{\xi}_n.$$

For each subsequence $\{n_i\}_{i=1}^{\infty}$ such that (ii)

$$\boldsymbol{\xi}_{n_i} \xrightarrow{\mathscr{D}} \boldsymbol{\xi}_0$$

for some random vector/matrix $\boldsymbol{\xi}_0$,

$$\boldsymbol{\xi}_{n_i}^* \big| \boldsymbol{\zeta}_{n_i} \xrightarrow{\mathscr{D}(\mathbb{P})} \boldsymbol{\xi}_0$$

too.

(iii) For each subsequence $\{n_i\}_{i=1}^{\infty}$ there exists a subsequence $\{n_{i_k}\}_{k=1}^{\infty}$ such that $\boldsymbol{\xi}_{n_{i_k}}^*$ conditional on $\zeta_{n_{i_{\nu}}}$ converges in distribution in probability \mathbb{P} to the distributional limit of $\xi_{n_{i_{\nu}}}$ as $k \to \infty$.

We need to extend Slutsky's theorem for our "bootstrap world", i.e., to have a stability property for conditional distributions.

Theorem 1. Slutsky's extended theorem. Suppose that $\{\xi_n^*, \zeta_n^*, \chi_n\}_{n=1}^{\infty}$ are sequences of random vectors/matrices. Then,

$$\boldsymbol{\xi}_{n}^{*} | \boldsymbol{\chi}_{n} \xrightarrow[n \to \infty]{\mathscr{D}([\mathbb{P}] - a.s.)} \boldsymbol{\xi}_{0}$$
(3)

and

$$\zeta_n^* \xrightarrow[n \to \infty]{\mathbb{P}^*_{\chi_n}([\mathbb{P}]^{-a.s.})} \zeta_0, \tag{4}$$

where ξ_0 is a random matrix/vector and ζ_0 is a non-random element, implies (for suitable vector/matrix dimensions):

(i) $[\boldsymbol{\xi}_n^*, \boldsymbol{\zeta}_n^*] | \boldsymbol{\chi}_n \xrightarrow{\mathscr{D}([\mathbb{P}]^{-a.s.})}{n \to \infty} [\boldsymbol{\xi}_0, \boldsymbol{\zeta}_0];$ (ii) $[\boldsymbol{\zeta}_n^*, \boldsymbol{\xi}_n^*] | \boldsymbol{\chi}_n \xrightarrow{\mathscr{D}([\mathbb{P}]^{-a.s.})}{[\boldsymbol{\zeta}_n^*, \boldsymbol{\xi}_n^*]} [\boldsymbol{\zeta}_n^*, \boldsymbol{\xi}_n^*];$

(ii)
$$[\xi_n, \xi_n] | \chi_n \xrightarrow[n \to \infty]{n \to \infty} [\xi_0, \xi_0],$$

(iii) $\xi_n^* + \zeta_n^* | \chi_n \xrightarrow{\mathscr{D}([\mathbb{P}] - a.s.)} \xi_0 + \zeta_0,$

(iv) $\xi_n^* \zeta_n^* | \chi_n \xrightarrow{\mathscr{D}([\mathbb{P}]^{-a.s.})}_{n \to \infty} \xi_0 \zeta_0;$

(v)
$$\zeta_n^* \xi_n^* | \chi_n \xrightarrow{\mathscr{D}([\mathbb{P}] - a.s.)}_{n \to \infty} \zeta_0 \xi_0;$$

- (vi) $(\zeta_n^*)^{-1} \xi_n^* | \chi_n \xrightarrow{\mathscr{D}([\mathbb{P}]^{-a.s.})}{\xrightarrow{n \to \infty}} \zeta_0^{-1} \xi_0$, provided that ζ_n^* and ζ_0 are invertible; (vii) $\xi_n^*(\zeta_n^*)^{-1} | \chi_n \xrightarrow{\mathscr{D}([\mathbb{P}]^{-a.s.})}{\xrightarrow{n \to \infty}} \xi_0 \zeta_0^{-1}$, provided that ζ_n^* and ζ_0 are invertible.
- Moreover,

$$\boldsymbol{\xi}_{n}^{*} | \boldsymbol{\chi}_{n} \xrightarrow[n \to \infty]{\mathscr{D}(\mathbb{P})} \boldsymbol{\xi}_{0}$$

$$\tag{5}$$

and

$$\zeta_n^* \xrightarrow[n \to \infty]{\mathbb{P}_{\chi_n}^*(\mathbb{P})} \zeta_0, \tag{6}$$

where ξ_0 is a random matrix/vector and ζ_0 is a non-random element, implies (for suitable vector/matrix dimensions):

$$\begin{array}{ll} (viii) & \left[\boldsymbol{\xi}_{n}^{*},\boldsymbol{\zeta}_{n}^{*}\right] \middle| \boldsymbol{\chi}_{n} \xrightarrow{\mathscr{D}(\mathbb{P})} & \left[\boldsymbol{\xi}_{0},\boldsymbol{\zeta}_{0}\right]; \\ (ix) & \left[\boldsymbol{\zeta}_{n}^{*},\boldsymbol{\xi}_{n}^{*}\right] \middle| \boldsymbol{\chi}_{n} \xrightarrow{\mathscr{D}(\mathbb{P})} & \left[\boldsymbol{\zeta}_{0},\boldsymbol{\xi}_{0}\right]; \\ (x) & \boldsymbol{\xi}_{n}^{*} + \boldsymbol{\zeta}_{n}^{*} \middle| \boldsymbol{\chi}_{n} \xrightarrow{\mathscr{D}(\mathbb{P})} & \boldsymbol{\xi}_{0} + \boldsymbol{\zeta}_{0}; \\ (vi) & \mathbf{\chi}_{n}^{*} \mathbf{\chi}_{n} \xrightarrow{\mathscr{D}(\mathbb{P})} & \mathbf{\chi}_{n} \mathbf{\chi}_{n} \end{array}$$

$$\begin{array}{ccc} (\chi l) & \zeta_n \zeta_n | \chi_n \xrightarrow[n \to \infty]{} \zeta_0 \zeta_0; \\ (\chi l) & \zeta^* \chi^* | \chi \xrightarrow{\mathcal{D}} (\mathbb{P}) & \zeta \chi \chi \end{array}$$

(xii) $\zeta_n^* \xi_n^* | \chi_n \xrightarrow[n \to \infty]{} \zeta_0 \zeta_0;$ (xiii) $(\zeta_n^*)^{-1} \xi_n^* | \chi_n \xrightarrow[n \to \infty]{} \zeta_0^{-1} \xi_0, \text{ provided that } \zeta_n^* \text{ and } \zeta_0 \text{ are invertible;}$ (xiv) $\xi_n^* (\zeta_n^*)^{-1} | \chi_n \xrightarrow[n \to \infty]{} \xi_0 \zeta_0^{-1}, \text{ provided that } \zeta_n^* \text{ and } \zeta_0 \text{ are invertible.}$

3.2. Weak Dependence

In order to overcome the assumption of independent observations, the dependence between data needs to be specified. It is assumed that $\{\xi_n\}_{n=1}^{\infty}$ is a sequence of random elements on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. For sub- σ -fields $\mathcal{A}, \mathcal{B} \subseteq \mathcal{F}$, we define

$$\alpha(\mathcal{A},\mathcal{B}) := \sup_{A \in \mathcal{A}, B \in \mathcal{B}} |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|,$$
$$\varphi(\mathcal{A},\mathcal{B}) := \sup_{A \in \mathcal{A}, B \in \mathcal{B}, \mathbb{P}(A) > 0} |\mathbb{P}(B|A) - \mathbb{P}(B)|.$$

Intuitively, α and φ measure the dependence of the events in \mathcal{B} on those in \mathcal{A} . Henceforth, let us define a filtration $\mathcal{F}_m^n := \sigma(\xi_i \in \mathcal{F}, m \leq i \leq n)$.

There are many ways how to describe weak dependence or, in other words, asymptotic independence of random variables [23]. Here, we concentrate on two approaches, however other ways of defining dependence can be involved, [24]. A sequence $\{\xi_n\}_{n=1}^{\infty}$ of random elements (e.g., variables) is said to be *strong mixing* (α -mixing) if

$$\alpha(n) := \sup_{k \in \mathbb{N}} \alpha(\mathcal{F}_1^k, \mathcal{F}_{k+n}^\infty) \to 0, \quad n \to \infty;$$
(7)

Moreover, it is said to be *uniformly strong mixing* (φ -mixing) if

$$\varphi(n) := \sup_{k \in \mathbb{N}} \varphi(\mathcal{F}_1^k, \mathcal{F}_{k+n}^\infty) \to 0, \quad n \to \infty.$$
(8)

Uniformly strong mixing–introduced by [25]–implies strong mixing [26], which was presented by [27]. Coefficients of dependence $\alpha(n)$ and $\varphi(n)$ measure how much dependence exists between events separated by at least *n* observations or time periods [28].

In [29], a class of *m*-dependent processes was comprehensively and intensively analyzed. These types of time series are φ -mixing, since are finite order ARMA processes with innovations satisfying *Doeblin's condition*, ([30], p. 168) or ([31], p. 192). Hence, the ARMA processes with continuously distributed stationary innovations and bounded variance are α -mixing and, thus, φ -mixing. Finite order processes, which do not satisfy Doeblin's condition, can be shown to be α -mixing ([32], pp. 312–313). Reference [33] provides general conditions under which stationary Markov processes are α -mixing. Since functions of mixing processes are themselves mixing [23], time-varying functions of any of the processes just mentioned are mixing as well.

It has to be emphasized that any form of errors' *stationarity* is not assumed. Omitting this, sometimes restrictive, assumption strengthen our results. It is obvious that $\alpha(\mathcal{A}, \mathcal{B}) = \alpha(\mathcal{B}, \mathcal{A})$ for arbitrary sub- σ -fields $\mathcal{A}, \mathcal{B} \subseteq \mathcal{F}$. This type of symmetry does not hold for φ -dependence. Indeed, ([33], pp. 213–214) constructed some strictly stationary Markov chains that are

 φ -mixing but not "time-reversed" φ -mixing. Therefore, it is not possible to "interchange" the past with the future regarding the definition of the φ -mixing coefficient.

A strong law of large numbers (SLLN) for α -dependent non-identically distributed variables needs to be recalled.

Lemma 2. Strong law of large numbers for α -mixing. Let $\{\xi_n\}_{n=1}^{\infty}$ be a sequence of α -mixing random variables satisfying

$$\sup_{n \in \mathbb{N}} \mathbb{E}|\xi_n|^q < \infty \tag{9}$$

for some q > 1. Suppose that there exists $\delta > 0$ such that as $n \to \infty$,

$$\alpha(n) = \begin{cases} \mathcal{O}\left(n^{-\frac{q}{2q-2}-\delta}\right) & \text{if } 1 < q < 2, \\ \mathcal{O}\left(n^{-\frac{2}{q}-\delta}\right) & \text{if } q \ge 2. \end{cases}$$
(10)

then

$$\lim_{n \to \infty} \frac{\sum_{i=1}^{n} (\xi_i - \mathbb{E}\xi_i)}{n} = 0 \quad a.s$$

Furthermore, a SLLN for φ -dependent non-identically distributed variables is desired as well.

Lemma 3. Strong law of large numbers for φ -mixing. Let $\{\xi_n\}_{n=1}^{\infty}$ be a sequence of zero mean φ -mixing random variables satisfying

$$\sum_{n=1}^{\infty} \sqrt{\varphi(n)} < \infty \tag{11}$$

and let $\{b_n\}_{n=1}^{\infty}$ be a non-decreasing unbounded sequence of positive numbers. Assume that

$$\sum_{n=1}^{\infty} \frac{\mathbb{E}\xi_n^2}{b_n^2} < \infty, \tag{12}$$

then

$$\lim_{n\to\infty}\frac{\sum_{i=1}^n\xi_i}{b_n}=0\quad a.s.$$

For a given random sequence $\xi_{\circ} \equiv {\xi_n}_{n=1}^{\infty}$ of random elements, the dependence coefficients $\alpha(n)$ will be denoted $\alpha(\xi_{\circ}, n)$. Analogous notation is used for φ -mixing sequences. Moreover, an auxiliary lemma for latter application of the SLLN for non-identically distributed random variables is stated.

The following lemma describes an asymptotic behavior of α - and φ -mixing coefficients of the corresponding random sequences after a transformation. More precisely, the Borel transformation preserves the property of α - and φ -mixing and, moreover, sustains the rate of the mixing coefficients.

Lemma 4. Suppose that for each $m = 1, 2, ..., \xi_{\circ}^{(m)} := \{\xi_k^{(m)}\}_{k \in \mathbb{Z}}$ is a sequence of random variables. Suppose the sequences $\xi_{\circ}^{(m)}$, m = 1, 2, ... are independent of each other. Suppose that for each $k \in \mathbb{Z}$, $h_k : \mathbb{R} \times \mathbb{R} \times ... \to \mathbb{R}$ is a Borel function. Define the sequence $\xi_{\circ} := \{\xi_k\}_{k \in \mathbb{Z}}$ of random variables by

$$\xi_k := h_k \Big(\xi_k^{(1)}, \xi_k^{(2)}, \ldots \Big), \quad k \in \mathbb{Z}.$$

then for each $n \ge 1$, the following statements hold:

- (i) $\alpha(\xi_{\circ}, n) \leq \sum_{m=1}^{\infty} \alpha(\xi_{\circ}^{(m)}, n),$
- (ii) $\varphi(\xi_{\circ}, n) \leq \sum_{m=1}^{\infty} \varphi(\xi_{\circ}^{(m)}, n)$

Let $S_n := \sum_{i=1}^n \xi_i$ and $\zeta_n^2 := \mathbb{V}$ ar S_n . For a random element from the Skorokhod space D[0, 1],

$$\mathcal{W}_n(t) := \frac{S_{[nt]}}{\varsigma_n}, \quad 0 \le t \le 1,$$
(13)

where $[\cdot]$ denotes the nearest integer function, a *functional central limit theorem*—also called a *weak invariance principle*—can be applied. This principle, now for α -mixing variables, will be postulated.

Lemma 5. WIP for α -mixing. Let $\{\xi_n\}_{n=1}^{\infty}$ be a sequence of zero mean α -mixing random variables with

$$\sup_{n\in\mathbb{N}}\mathbb{E}|\xi_n|^{2+\omega}<\infty\tag{14}$$

and

$$\sum_{n=1}^{\infty} \alpha(n)^{\omega/(2+\omega)} < \infty \tag{15}$$

for some $\omega > 0$. Suppose that

$$\frac{\mathbb{E}S_n^2}{n} \to \varsigma^2 > 0, \quad n \to \infty$$
(16)

is satisfied. Then

$$\mathcal{W}_n \xrightarrow{D[0,1]} \mathcal{W}, \quad n \to \infty.$$
 (17)

where W stands for the standard Wiener process.

Since the central limit theorem is just a special case of the weak invariance principle, then a corollary of previous Lemma 5 can be stated.

Corollary 1. Central limit theorem for α -mixing. Suppose that all the assumptions of Lemma 5 on a sequence of zero mean α -mixing random variables $\{\xi_n\}_{n=1}^{\infty}$ are satisfied. Then

$$\frac{S_n}{\varsigma_n} \xrightarrow{\mathscr{D}} \mathcal{N}(0,1), \quad n \to \infty.$$
(18)

Lemma 6. Lindeberg central limit theorem for φ -mixing. Let $\{\xi_n\}_{n=1}^{\infty}$ be a sequence of zero mean φ -mixing random variables having finite variance. Suppose that the Lindeberg condition

$$\forall \delta > 0: \quad \lim_{n \to \infty} \frac{1}{\varsigma_n^2} \sum_{i=1}^n \mathbb{E} \xi_i^2 \mathcal{I}\{|\xi_i| > \delta \varsigma_n\} = 0 \tag{19}$$

is satisfied. Then

$$\frac{S_n}{\varsigma_n} \xrightarrow{\mathscr{D}} \mathcal{N}(0,1), \quad n \to \infty.$$

The Lindeberg condition (19) can be replaced by a stronger type of the Lyapunov condition. This fact leads into the following corollary, which is more comfortable for us from the point of applicability.

Corollary 2. Central limit theorem for φ -mixing. Let $\{\xi_n\}_{n=1}^{\infty}$ be a sequence of zero mean φ -mixing random variables such that

$$\sup_{n\in\mathbb{N}}\mathbb{E}|\xi_n|^{2+\omega}<\infty\tag{20}$$

for some $\omega > 0$ *and*

$$\frac{\mathbb{E}S_n^2}{n} \to \varsigma^2 > 0, \quad n \to \infty.$$
(21)

then

$$\frac{S_n}{S_n} \xrightarrow{\mathscr{D}} \mathcal{N}(0,1), \quad n \to \infty.$$
(22)

Assumption (21) may even be replaced by a weaker one:

$$\liminf_{n\to\infty}\frac{\mathbb{E}S_n^2}{n}=\varsigma^2>0,$$

where the limit inferior is used instead of the original limit. Assuming that a sequence of random variables is φ -mixing implies that this sequence is α -mixing. On the other hand, the central limit theorem for φ -mixing (Corollary 2) has weaker assumptions than the central limit theorem for α -mixing (Corollary 1). Indeed, Corollary 2 does not require any assumption on *mixing rate* $\varphi(n)$ such as Assumption (15) on α -mixing rates.

The previous theoretical results contain the variance ζ_n^2 for the underlying process of random variables. One may additionally assume that

$$\lim_{n\to\infty}\frac{\zeta_n^2}{n}=\zeta^2,$$

where ζ^2 is the so-called *long run variance*, which sometimes needs to be estimated. If $\{\xi_n\}_{n=1}^{\infty}$ are zero mean and stationary, then this long run variance can be decomposed in the following way

$$\varsigma^2 = \mathbb{E}\xi_1^2 + 2\sum_{i=1}^{\infty} \mathbb{E}\xi_1 \xi_{i+1}.$$

Often, the Bartlett estimator is used to estimate the long run variance, i.e.,

$$\hat{\varsigma}_n^2(m) = \hat{R}(0) + 2 \sum_{1 \le k \le m} \left(1 - \frac{k}{m}\right) \hat{R}(k), \quad m < n,$$

where

$$\hat{R}(k) = \frac{1}{n} \sum_{1 \le i \le n-k} (\tilde{\xi}_i - \bar{\xi}_n) (\tilde{\xi}_{i+k} - \bar{\xi}_n), \quad 0 \le k < n$$

The consistency properties of the above described Bartlett estimator and of its modification are studied in [34]. Other similar types of estimators can be used instead, for instance Parzen kernels [35].

4. Bootstrap Laws of Large Numbers

A theoretical mid-step for showing validity of the bootstrap methods are the bootstrap laws of large numbers. From now on, we mean by a *bootstrap version* of $\boldsymbol{\xi} \equiv [\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_n]^\top$ its (randomly) resampled sequence with replacement—denoted by $\boldsymbol{\xi}^* \equiv [\boldsymbol{\xi}_1^*, \dots, \boldsymbol{\xi}_n^*]^\top$ —with the same length, where for each $i \in \{1, \dots, n\}$ holds $\mathbb{P}_{\boldsymbol{\xi}}^*[\boldsymbol{\xi}_i^* = \boldsymbol{\xi}_j] = 1/n, j = 1, \dots, n$. So, $\boldsymbol{\xi}_i^*$ has a discrete uniform distribution on $\{\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_n\}$ for every $i = 1, \dots, n$.

4.1. Bootstrap Weak LLN for Independent Data

First of all, a base stone for the consistency of the independent bootstrap lies in the *bootstrap weak law of large numbers* (BWLLN).

Theorem 2. Bootstrap weak law of large numbers. Let $\{\xi_n\}_{n=1}^{\infty}$ be a sequence of independent random variables. If

$$\sup_{n\in\mathbb{N}}\mathbb{E}_{\mathbb{P}}\xi_n^2<\infty,\tag{23}$$

then

$$n^{-1}\sum_{i=1}^{n}\xi_{i}^{*}-n^{-1}\sum_{i=1}^{n}\xi_{i}\xrightarrow{\mathbb{P}_{\xi}^{*}(\mathbb{P})}{\xrightarrow{n\to\infty}}0,$$
(24)

where $\boldsymbol{\xi}^* \equiv [\xi_1^*, \dots, \xi_n^*]^\top$ is the bootstrapped version of $\boldsymbol{\xi} \equiv [\xi_1, \dots, \xi_n]^\top$.

4.2. Bootstrap Weak LLNs for NINID

Now, the BWLLN for the non-stationary α -mixing is given.

Theorem 3. Bootstrap weak law of large numbers for α -mixing. Let $\{\xi_n\}_{n=1}^{\infty}$ be a sequence of zero mean α -mixing random variables satisfying

$$\sup_{n\in\mathbb{N}}\mathbb{E}\zeta_n^2<\infty.$$
(25)

assume that there exists $\delta > 0$ such that

$$\alpha(n) = \mathcal{O}\left(n^{-1-\delta}\right), \quad n \to \infty.$$
(26)

if $b \to \infty$ and $b = o(n^{1/2})$, $n \to \infty$, then, under the MBB Procedure 2,

$$n^{-1}\sum_{i=1}^n \xi_i^* - n^{-1}\sum_{i=1}^n \xi_i \xrightarrow{\mathbb{P}_{\xi}^*(\mathbb{P})} 0,$$

where $\boldsymbol{\xi}^* \equiv [\xi_1^*, \dots, \xi_n^*]^\top$ is the MBB version of $\boldsymbol{\xi} \equiv [\xi_1, \dots, \xi_n]^\top$.

Similarly, the BWLLN is stated for the φ -mixing sequences.

Theorem 4. Bootstrap weak law of large numbers for φ -mixing. Let $\{\xi_n\}_{n=1}^{\infty}$ be a sequence of zero mean φ -mixing random variables satisfying

$$\sum_{n=1}^{\infty} \frac{\mathbb{E}\xi_n^2}{n^2} < \infty$$

and

$$\sum_{n=1}^{\infty} \sqrt{\varphi(n)} < \infty.$$
(27)

if $b \to \infty$ *and* $b = o(n^{1/2})$ *,* $n \to \infty$ *, then, under the MBB Procedure 2,*

$$n^{-1}\sum_{i=1}^n \xi_i^* - n^{-1}\sum_{i=1}^n \xi_i \xrightarrow{\mathbb{P}_{\xi}^*(\mathbb{P})} 0,$$

where $\boldsymbol{\xi}^* \equiv [\xi_1^*, \dots, \xi_n^*]^\top$ is the MBB version of $\boldsymbol{\xi} \equiv [\xi_1, \dots, \xi_n]^\top$.

5. Bootstrap Central Limit Theorems

If a statistic converges to a normal distribution, then the interest is to asymptotically compare this original limiting distribution to the bootstrap one. The main theoretical results for the asymptotic validity of the bootstrap methods lie in showing that the bootstrap distribution properly mimics the asymptotic one.

5.1. Bootstrap CLT for Independent Data

A pylon for the proof of central limit theorem for the bootstrapped sample is created by an extension of the well-known Berry–Esseen theorem. **Theorem 5.** Berry–Esseen-Katz theorem. Let g be a non-negative, even, non-decreasing function on $[0, \infty)$ satisfying:

- (i) $\lim_{x\to\infty} g(x) = \infty$,
- (ii) x/g(x) is defined for all $x \in \mathbb{R}$ and non-decreasing on $[0, \infty)$.

Assume that $\{\xi_n\}_{n=1}^{\infty}$ are IID random variables such that $\mathbb{E}\xi_1 = 0$ and $\mathbb{V}ar \xi_1 = \zeta^2 > 0$. If

$$\mathbb{E}\xi_1^2 g(\xi_1) < \infty,$$

then there exists a constant C > 0, such that for all $n \in \mathbb{N}$ holds

$$\sup_{x\in\mathbb{R}}\left|\mathbb{P}\left[\frac{1}{\sqrt{n\zeta^2}}\sum_{i=1}^n\zeta_i\leq x\right]-\int_{-\infty}^x\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{t^2}{2}\right\}dt\right|\leq\frac{C\mathbb{E}\zeta_1^2g(\zeta_1/\zeta)}{\zeta^2g(\sqrt{n})}.$$

A *bootstrap central limit theorem* provides the desired approximate distributional closeness and, thus, appropriateness of the bootstrap.

Theorem 6. Bootstrap central limit theorem for independent variables. Let $\{\xi_n\}_{n=1}^{\infty}$ be a sequence of zero mean independent random variables satisfying

$$\sup_{n\in\mathbb{N}}\mathbb{E}_{\mathbb{P}}\xi_{n}^{4}<\infty. \tag{28}$$

suppose that $\boldsymbol{\xi}^* \equiv [\boldsymbol{\xi}_1^*, \dots, \boldsymbol{\xi}_n^*]^\top$ is the bootstrapped version of $\boldsymbol{\xi} \equiv [\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_n]^\top$ and denote

$$\bar{\xi}_n := n^{-1} \sum_{i=1}^n \xi_i, \quad \bar{\xi}_n^* := n^{-1} \sum_{i=1}^n \xi_i^*, \quad and \quad \zeta_n^2 := \sum_{i=1}^n \mathbb{V}ar_{\mathbb{P}}\xi_i.$$

if

$$\liminf_{n \to \infty} \frac{\zeta_n^2}{n} = \zeta^2 > 0, \tag{29}$$

then

$$\sup_{x \in \mathbb{R}} \left| \mathbb{P}_{\xi}^{*} \left[\frac{n}{\sqrt{\zeta_{n}^{2}}} (\bar{\zeta}_{n}^{*} - \bar{\zeta}_{n}) \leq x \right] - \mathbb{P} \left[\frac{n}{\sqrt{\zeta_{n}^{2}}} \bar{\zeta}_{n} \leq x \right] \right| \xrightarrow[n \to \infty]{} 0.$$
(30)

Assumption (28) may seem a little bit restrictive. It may be weakened to

$$\sum_{n=1}^{\infty} \frac{\mathbb{E}_{\mathbb{P}} \xi_n^4}{n^2} < \infty \quad \text{and} \quad \sup_{n \in \mathbb{N}} \mathbb{E}_{\mathbb{P}} |\xi_n|^{2+\omega} < \infty, \text{ for some } \omega > 0;$$

As found out while going through the proof of Theorem 6. On the other hand, the previous weaker premises are rather complicated to verify for transformed errors in the proof of bootstrap consistency, which will be noticed later on. Reference [36] strengthened Assumption (28) and replaced by

$$\sup_{n\in\mathbb{N}}\mathbb{E}_{\mathbb{P}}|\sqrt{n}\xi_{n}|^{4+\omega}<\infty,\tag{31}$$

For some $\omega > 0$. Afterwards, bootstrap CLT 6 provides a stronger result:

$$\sup_{x \in \mathbb{R}} \left| \mathbb{P}_{\xi}^{*} \left[\frac{n}{\sqrt{\varsigma_{n}^{2}}} \left(\bar{\xi}_{n}^{*} - \bar{\xi}_{n} \right) \leq x \right] - \mathbb{P} \left[\frac{n}{\sqrt{\varsigma_{n}^{2}}} \bar{\xi}_{n} \leq x \right] \right| \xrightarrow[n \to \infty]{[\mathbb{P}]^{-a.s.}} 0, \tag{32}$$

whereas the convergence in distribution in probability is replaced by the convergence in distribution almost surely. In spite of this, Assumption (31) can be considered as too much restrictive.

Our situation would become much easier, when IID variables are assumed. Let us have a look at the proof of Theorem 6. If $\mathbb{E}_{\mathbb{P}}|\xi_1|^{2+\omega} < \infty$ for some $\omega > 0$ is additionally assumed, then the right-hand side of (A7) converges to zero $[\mathbb{P}]$ -almost surely. In addition, finite $(2 + \omega)$ -th moment would be enough to prove (*ii*) and, thus, Relation (32) holds under such conditions.

The *equiboundedness of the fourth moments* (28) in the bootstrap CLT is needed, because the second conditional moment is necessary for the existence of $\operatorname{Var}_{\mathbb{P}^*_{\xi}}\xi_1^*$ and, consequently, the equiboundedness of the second moment of the second conditional moment is used for the convergence $[\mathbb{P}]$ -almost surely of $\operatorname{Var}_{\mathbb{P}^*_{\xi}}\xi_1^*$. The equiboundedness of the fourth moments in the BCLT can be weakened when bearing in mind identically distributed variables.

Let us concentrate on the conditional variance of sum of the resampled data. The conditional variance \mathbb{V} ar $\mathbb{P}^*_{\xi} \xi_1^*$ used for the normalization in (30) may be simply replaced by $n^{-1} \zeta_n^2$, see the proof of the BCLT (6). Therefore, is makes the assertion of the BCLT even stronger and, hence, more applicable.

A utilization of the Cramér–Wold device helps us to derive a bootstrap version of the CLT for random vectors.

Theorem 7. Bootstrap multivariate central limit theorem for independent vectors. Let $\{\xi_n\}_{n=1}^{\infty}$ be a sequence of zero mean independent q-dimensional random vectors satisfying

$$\sup_{n \in \mathbb{N}} \mathbb{E}_{\mathbb{P}} |\xi_{j,n}|^4 < \infty, \quad j \in \{1, \dots, q\},$$
(33)

where $\boldsymbol{\xi}_n \equiv [\xi_{1,n}, \dots, \xi_{q,n}]^\top \in \mathbb{R}^q$, $n \in \mathbb{N}$. Assume that $\boldsymbol{\Xi}^* \equiv [\boldsymbol{\xi}_1^*, \dots, \boldsymbol{\xi}_n^*]^\top$ is the bootstrapped version of $\boldsymbol{\Xi} \equiv [\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_n]^\top$. Denote

$$\bar{\xi}_n := n^{-1} \sum_{i=1}^n \xi_i, \quad \bar{\xi}_n^* := n^{-1} \sum_{i=1}^n \xi_i^*, \quad and \quad \Gamma_n := \sum_{i=1}^n \mathbb{V}ar_{\mathbb{P}}\xi_i.$$

if

$$\lim_{n\to\infty}\frac{1}{n}\Gamma_n=\Gamma>\mathbf{0},\tag{34}$$

then

$$n\Gamma_n^{-1/2} \left(\bar{\xi}_n^* - \bar{\xi}_n \right) \left| \Xi \xleftarrow{\mathscr{D}(\mathbb{P})}_{n \to \infty} n\Gamma_n^{-1/2} \bar{\xi}_n \right. \tag{35}$$

and, moreover,

$$\sqrt{n} \left(\bar{\xi}_n^* - \mathbb{E}_{\mathbb{P}^*} \xi_1^* \right) \Big| \Xi \xleftarrow{\mathscr{D}(\mathbb{P})}_{n \to \infty} \sqrt{n} \bar{\xi}_n.$$
(36)

5.2. Bootstrap CLTs for NINID

Lastly, the central limit theorems for the bootstrapped sample mean from *non-stationary strong mixing or uniformly strong mixing* sequences are stated.

Theorem 8. Bootstrap central limit theorem for α -mixing. Let $\{\xi_n\}_{n=1}^{\infty}$ be a sequence of zero mean α -mixing random variables with

$$\sup_{n\in\mathbb{N}}\mathbb{E}|\xi_n|^{4+\omega}<\infty\tag{37}$$

and

$$\alpha(n) = \mathcal{O}(n^{-1-\delta}), \quad n \to \infty$$
 (38)

for some $\omega > 0$ and $\delta > 0$ such that

$$\frac{4}{\omega} < \delta. \tag{39}$$

denote

$$\bar{\xi}_n := n^{-1} \sum_{i=1}^n \xi_i \quad and \quad \bar{\xi}_n^* := n^{-1} \sum_{i=1}^n \xi_i^*.$$

suppose that

$$\liminf_{n \to \infty} \frac{\mathbb{E}S_n^2}{n} = \varsigma^2 > 0 \tag{40}$$

is satisfied. If $b \to \infty$ *and* $b = o(n^{1/2})$ *,* $n \to \infty$ *, then, under the MBB Procedure 2,*

$$\sup_{x \in \mathbb{R}} \left| \mathbb{P}_{\xi}^{*} \left(\frac{n}{\sqrt{\mathbb{E}S_{n}^{2}}} (\bar{\xi}_{n}^{*} - \bar{\xi}_{n}) \leq x \right) - \mathbb{P} \left(\frac{n}{\sqrt{\mathbb{E}S_{n}^{2}}} \bar{\xi}_{n} \leq x \right) \right| \xrightarrow{\mathbb{P}} 0, \quad n \to \infty,$$
(41)

where $\boldsymbol{\xi}^* \equiv [\xi_1^*, \dots, \xi_n^*]^\top$ is the MBB version of $\boldsymbol{\xi} \equiv [\xi_1, \dots, \xi_n]^\top$.

Similar theorem for IID random variables was proved by [37]. Furthermore, a version of this BCLT in case of stationary α -mixing is given by [11].

Theorem 9. Bootstrap central limit theorem for φ -mixing. Let $\{\xi_n\}_{n=1}^{\infty}$ be a sequence of zero mean φ -mixing random variables with

$$\sup_{n\in\mathbb{N}}\mathbb{E}|\xi_n|^4<\infty\tag{42}$$

and

$$\varphi(n) = \mathcal{O}(n^{-2-\delta}), \quad n \to \infty$$
 (43)

for some $\delta > 0$. Denote

$$\bar{\xi}_n := n^{-1} \sum_{i=1}^n \xi_i \quad and \quad \bar{\xi}_n^* := n^{-1} \sum_{i=1}^n \xi_i^*.$$

suppose that

$$\liminf_{n \to \infty} \frac{\mathbb{E}S_n^2}{n} = \zeta^2 > 0 \tag{44}$$

is satisfied. If $b \to \infty$ and $b = o(n^{1/2})$, $n \to \infty$, then, under the MBB Procedure 2,

$$\sup_{x \in \mathbb{R}} \left| \mathbb{P}_{\xi}^{*} \left(\frac{n}{\sqrt{\mathbb{E}S_{n}^{2}}} (\bar{\xi}_{n}^{*} - \bar{\xi}_{n}) \leq x \right) - \mathbb{P} \left(\frac{n}{\sqrt{\mathbb{E}S_{n}^{2}}} \bar{\xi}_{n} \leq x \right) \right| \xrightarrow{\mathbb{P}} 0, \quad n \to \infty,$$
(45)

where $\boldsymbol{\xi}^* \equiv [\xi_1^*, \dots, \xi_n^*]^\top$ is the MBB version of $\boldsymbol{\xi} \equiv [\xi_1, \dots, \xi_n]^\top$.

With respect to Theorem 7, one may also postulate the multivariate version of the central limit theorem for the α - and φ -mixing accordingly, based on the bootstrap univariate CLT (8) and (9) using the Cramér–Wold device.

6. Real Data Analyses

If the data are indeed not independent and/or not identically distributed, using a proper resampling method can lead to, for instance, *improvement in precision*. We are going to deal with two real problems, where the first one comes from *psychometric evaluation* and the second one has grounds in *non-life insurance*. Handling the data as NINID will provide narrower confidence intervals for the mean parameter in both cases compared to the incorrect approach, when the data are falsely considered as IID.

6.1. Psychometry

Psychometric evaluation through a test with many binary (i.e., 0–1) answers from one subject naturally brings a question, what is the probability of the "correct" answer. This parameter of interest for the underlying Bernoulli distribution is nothing else than its

mean. Thus, the sample mean is the consistent and asymptotically normal estimator of the parameter of probability of the correct answer. If the data are dependent, one cannot rely on normal asymptotics when constructing the confidence interval for the mean of the Bernoulli distribution, because the corresponding limiting variance is unknown [38–40]. Therefore, bootstrapping can provide a suitable solution. However, a proper bootstrap technique needs to be chosen and applied.

The above described issue is going to be illustrated on the dataset from "Programme for International Student Assessment (PISA) 2012 U.S. Math Assessment", which can be downloaded from (https://tmsalab.github.io/edmdata/, accessed on 12 November 2022). This dataset was discussed in [41]. We concentrated on one subject, who was randomly selected (ID = 15) and s/he responded on n = 34 items. The dependence between the subject's answers comes from the fact that all the items are answered by the same subject, having the same abilities, knowledge, and state of mind during the examination.

We utilize the independent bootstrap (ignoring the dependence among the responses) and the moving block bootstrap with the blocksize $b = \lfloor n^{1/3} \rfloor = 3$. The results for resampling the sample mean are shown in Table 1 and in Figure 1.

Table 1. Empirical distributional quantities of the sample mean based on the independent bootstrap (incorrect approach) against moving block bootstrap (proper approach) for the correct/incorrect answers by a subject from the Programme for International Student Assessment (PISA) 2012–U.S. Math Assessment; number of observations n = 34.

Empirical Quantity	Independent Bootstrap	Moving Block Bootstrap
2.5th percentile	0.2647	0.3125
First quartile	0.3824	0.4062
Median	0.4412	0.4375
Third quartile	0.5000	0.4999
97.5th percentile	0.6176	0.5938



Figure 1. Empirical distributions of the sample mean based on the independent bootstrap (green—incorrect approach) against moving block bootstrap (orange—proper approach) for the correct/incorrect answers by a subject from the Programme for International Student Assessment (PISA) 2012—U.S. Math Assessment.

One may notice, from Table 1 as well as from Figure 1, that the difference between the 97.5th percentile and the 2.5th percentile, which serves as a 95% confidence interval, is narrower in case of the moving block bootstrap approach. Indeed, the 95% confidence interval based on independent bootstrap is [0.2647, 0.6176], whereas the 95% confidence interval based on moving block bootstrap is [0.3125, 0.5938]. Hence, the lengths of the confidence intervals are 0.3529 and 0.2813 for the independent and the moving block bootstrap, respectively, which yields approximately 20% reduction in the confidence interval's length and, thus, an increase in estimation precision. In the end, the standard deviation for the independent bootstrap is 0.0851, whereas the standard deviation for the moving block bootstrap drops down to 0.0764, which also provides an increase in precision by taking into account the underlying dependence among the data.

6.2. Insurance

In non-life insurance, the claims are traditionally split into the attritional and large claims, see, for example, [42]. One of the most fundamental tasks in non-life insurance, conducted on regular basis, is risk reserving assessment analysis, which amounts to predict the overall loss reserves to cover possible future claims. In order to do that, one firstly needs to estimate stochastically the distributional characteristics of the historical claims [43]. Here, we concentrate on estimation of the unknown mean for a reinsurance caped layer of large claims. Since the considered reinsurance layer will be bounded, there is no issue of assuming that the expectation of the underlying claims belonging to this layer exists.

We use the large fire insurance claims in Denmark from Thursday 3rd January 1980 until Monday 31st December 1990. The data called "Danish Fire Insurance Claims" were supplied by Mette Rytgaard of Copenhagen Re. They were described in [44] and can be freely downloaded from (https://tmsalab.github.io/edmdata/, accessed on 12 November 2022). Note that these data form an irregular time series, which does not contradict the notion of weak dependence. On contrary, the concept of mixing is very suitable for the unequally spaced time series [45]. We think of the reinsurance caped layer covering the large claims starting from 20M DKK (Danish crowns) up to 100M DKK, which consists n = 33 fire claims. The parameter of interest for the underlying unknown distribution is the mean. Again, it exists due to the fact that the layer is bounded from below as well as from above. Hence, the sample mean is the consistent and asymptotically normal estimator of the mean parameter. If the data are dependent and non-stationary (and, thus, not identically distributed), one cannot rely on normal asymptotics when constructing the confidence interval for the mean, because the corresponding limiting variance becomes unknown [46–48]. Therefore, the proper bootstrap method can provide a suitable solution.

The independent bootstrap (ignoring the dependence and non-stationarity among the fire claims) and the moving block bootstrap with the blocksize $b = \lfloor n^{1/3} \rfloor = 3$ are applied. The results for resampling the sample mean are numerically displayed in Table 2. Additional to that, the empirical distributions of the sample mean based on the both resampling techniques together with the empirical distributional quantities from Table 2 are graphically visualized in Figure 2.

Table 2. Empirical distributional quantities of the sample mean based on the independent bootstrap (blue—incorrect approach) against moving block bootstrap (red—proper approach) for the Danish Fire Insurance Claims belonging to the reinsurance layer starting from 20M DKK (Danish crowns) up to 100M DKK; number of observations n = 33.

Empirical Quantity	Independent Bootstrap	Moving Block Bootstrap
2.5th percentile	27.94	29.39
First quartile	30.28	31.20
Median	31.64	32.20
Third quartile	33.08	33.29
97.5th percentile	35.90	35.38



Distribution of sample mean

Figure 2. Empirical distributions of the sample mean based on the independent bootstrap (blue incorrect approach) against moving block bootstrap (red—proper approach) for the Danish Fire Insurance Claims belonging to the reinsurance layer starting from 20M DKK (Danish crowns) up to 100M DKK.

Table 2 as well as Figure 2 clearly reveal that the 95% confidence interval, which is the difference between the 97.5th percentile and the 2.5th percentile, is narrower in case of the moving block bootstrap procedure. Truly, the 95% confidence interval based on independent bootstrap is [27.94, 35.90], whereas the 95% confidence interval based on moving block bootstrap is [29.39, 35.38]. Hence, the lengths of the confidence intervals are 7.96 and 5.99 for the independent and the moving block bootstrap, respectively, which gives approximately 25% reduction in the confidence interval's length. This provides an increase in estimation precision, which can be incorporated and employed in the reserving techniques for not independent and not identically distributed data, see [49–52]. On top of that, the standard deviation for the independent bootstrap is 2.054, whereas the standard deviation for the moving block bootstrap drops down to 1.521, which again confirms an increase in accuracy by assuming the underlying dependence in the data.

7. Conclusions and Discussion

Asymptotic normality of the estimators or test statistics might be computationally unattainable and, hence, it can fail in practical applications. The discussed disadvantageous asymptotic properties lead to usage of the distributional-free resampling methods even for the IID observations. On top of that, the bootstrap procedures for a class of not independent and not identically distributed data are given together with their theoretical validity. There are three main methodological contributions to the bootstrap inference developed in this paper: (i) asymptotic closeness of the unknown stochastic quantities and their bootstrap counterparts is properly mathematically formalized; (ii) laws of large numbers for the bootstrapped α -mixing and φ -mixing random variables and vectors are postulated and proved; (iii) central limit theorems for the bootstrapped NINID observations are provided, which serve as justifications of the popular computer intensive techniques. Lastly, the theoretically valid approaches are practically performed on the problems from psychometry and insurance.

The concept of weak dependence provides a very flexible framework and it allows to handle any autocorrelation structure for the observations, if the mixing conditions are satisfied. Furthermore, if the parameter of interest is not the mean, our developed methodology can still be useful. Suppose that the corresponding estimator of the unknown parameter can be somehow linearized and, thus, rewritten as a continuous functional of a mean of the random variables plus some remainder, which is negligible in probability. This can be achieved, for instance, through the stochastic Taylor expansion. Then, one may apply the derived bootstrap machinery on the mean from the continuous functional and, additionally, use the continuous mapping theorem to obtain validity for bootstrapping the parameter of interest.

Author Contributions: All the authors M.H., M.M., B.P. and M.P.–contributed equally. All authors have read and agreed to the published version of the manuscript.

Funding: The research of Martin Hrba, Matúš Maciak, and Michal Pešta was funded by the Czech Science Foundation project GAČR No. 21–13323S. The work of Barbora Peštová was supported by the Czech Science Foundation project GAČR No. 21-03658S.

Data Availability Statement: There are two datasets used in this paper and both are publicly available. The first dataset–Programme for International Student Assessment (PISA) 2012 U.S. Math Assessment–can be downloaded from https://tmsalab.github.io/edmdata/, accessed on 12 November 2022 [41]. The second dataset–Danish Fire Insurance Claims–can be downloaded from https://vincentarelbundock.github.io/, accessed on 12 November 2022 [44].

Acknowledgments: The authors would like to thank three anonymous referees for their constructive comments and valuable remarks that have significantly improved this paper.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

IID	Independent and identically distributed
NINID	Not independent and not identically distributed
LLN	Law of large numbers
SLLN	Strong law of large numbers
WLLN	Weak law of large numbers
BWLLN	Bootstrap weak law of large numbers
CLT	Central limit theorem
BCLT	Bootstrap central limit theorem

Appendix A. Proofs

Proof of Lemma 1. A simple generalization of [53] (Lemma 1), where conditional law of $\xi_n^* | \zeta_n$ instead of ξ_n^* is considered in the proof as proposed by [54] (Proof of Theorem 4.1).

Proof of Theorem 1. First, we show that

$$[\boldsymbol{\xi}_n^*, \boldsymbol{\zeta}_0] | \boldsymbol{\chi}_n \xleftarrow{\mathscr{D}([\mathbb{P}]^{-a.s.})}_{n \to \infty} [\boldsymbol{\xi}_0, \boldsymbol{\zeta}_0], \tag{A1}$$

i.e., for arbitrary bounded continuous function f holds

$$\mathbb{E}_{\mathbb{P}^*_{\chi_n}} f([\boldsymbol{\xi}^*_n, \boldsymbol{\zeta}_0]) - \mathbb{E}_{\mathbb{P}} f([\boldsymbol{\xi}_0, \boldsymbol{\zeta}_0]) \xrightarrow{[\mathbb{P}]^{-a.s.}}_{n \to \infty} 0.$$
(A2)

Let $f([\cdot, \cdot])$ be such arbitrary bounded continuous function. Now consider the function of a single argument $g(\cdot) := f([\cdot, \zeta_0])$. This will obviously be a bounded and continuous non-random function as well. By assumption (3), we will have that

$$\mathbb{E}_{\mathbb{P}^*_{\boldsymbol{\chi}n}}g(\boldsymbol{\xi}^*_n) - \mathbb{E}_{\mathbb{P}}g(\boldsymbol{\xi}_n) \xrightarrow[n \to \infty]{[\mathbb{P}]^{-a.s.}} 0.$$

However, the latter expression is equivalent to (A2). Therefore, we now know that $[\xi_n^*, \zeta_0]$ conditioned on χ_n and $[\xi_0, \zeta_0]$ approach each other in distribution $[\mathbb{P}]$ -almost surely along χ_n .

Secondly, consider $\|[\xi_n^*, \zeta_n^*] - [\xi_n^*, \zeta_0]\| = \|\zeta_n^* - \zeta_0\|$. This expression converges in probability $\mathbb{P}_{\chi_n}^*$ to zero $[\mathbb{P}]$ -almost surely due to Assumption (4). Thus, we have demonstrated two facts: (A1) and

$$[\boldsymbol{\xi}_n^*,\boldsymbol{\zeta}_n^*] - [\boldsymbol{\xi}_n^*,\boldsymbol{\zeta}_0] \xrightarrow{\mathbb{P}_{\boldsymbol{\chi}_n}^*([\mathbb{P}]\text{-}a.s.)}{n \to \infty} \mathbf{0}.$$
(A3)

Since each bounded continuous function is bounded Lipschitz and vice versa, consider any bounded Lipschitz function $h(\cdot, \cdot)$:

$$\exists K, M > 0, \quad \forall \mathbf{x}_1, \mathbf{x}_2, \mathbf{y}_1, \mathbf{y}_2 : \\ h([\mathbf{x}_1, \mathbf{x}_2]) \le M \quad \& \quad |h([\mathbf{x}_1, \mathbf{x}_2]) - h([\mathbf{y}_1, \mathbf{y}_2])| \le K ||[\mathbf{x}_1, \mathbf{x}_2] - [\mathbf{y}_1, \mathbf{y}_2]||.$$

Take some arbitrary $\epsilon > 0$ and majorize

$$\begin{split} |\mathbb{E}_{\mathbb{P}_{\chi_{n}}^{*}}h([\xi_{n}^{*},\zeta_{n}^{*}]) - \mathbb{E}_{\mathbb{P}_{\chi_{n}}^{*}}h([\xi_{n}^{*},\zeta_{0}])| &\leq \mathbb{E}_{\mathbb{P}_{\chi_{n}}^{*}}|h([\xi_{n}^{*},\zeta_{n}^{*}]) - h([\xi_{n}^{*},\zeta_{0}])| \\ &= \mathbb{E}_{\mathbb{P}_{\chi_{n}}^{*}}[|h([\xi_{n}^{*},\zeta_{n}^{*}]) - h([\xi_{n}^{*},\zeta_{0}])|\mathcal{I}\{\|[\xi_{n}^{*},\zeta_{n}^{*}] - [\xi_{n}^{*},\zeta_{0}]\| &< \epsilon\}] \\ &+ \mathbb{E}_{\mathbb{P}_{\chi_{n}}^{*}}[|h([\xi_{n}^{*},\zeta_{n}^{*}]) - h([\xi_{n}^{*},\zeta_{0}])|\mathcal{I}\{\|[\xi_{n}^{*},\zeta_{n}^{*}] - [\xi_{n}^{*},\zeta_{0}]\| &\geq \epsilon\}] \\ &\leq \mathbb{E}_{\mathbb{P}_{\chi_{n}}^{*}}[K\|[\xi_{n}^{*},\zeta_{n}^{*}] - [\xi_{n}^{*},\zeta_{0}]\|\mathcal{I}\{\|[\xi_{n}^{*},\zeta_{n}^{*}] - [\xi_{n}^{*},\zeta_{0}]\| &< \epsilon\}] \\ &+ \mathbb{E}_{\mathbb{P}_{\chi_{n}}^{*}}[2M\mathcal{I}\{\|[\xi_{n}^{*},\zeta_{n}^{*}] - [\xi_{n}^{*},\zeta_{0}]\| &\geq \epsilon\}] \\ &\leq K\epsilon \mathbb{P}_{\chi_{n}}^{*}[\|[\xi_{n}^{*},\zeta_{n}^{*}] - [\xi_{n}^{*},\zeta_{0}]\| &< \epsilon\} + 2M\mathbb{P}_{\chi_{n}}^{*}[\|[\xi_{n}^{*},\zeta_{n}^{*}] - [\xi_{n}^{*},\zeta_{0}]\| &\geq \epsilon] \\ &\leq K\epsilon + 2M\mathbb{P}_{\chi_{n}}^{*}[\|[\xi_{n}^{*},\zeta_{n}^{*}] - [\xi_{n}^{*},\zeta_{0}]\| &\geq \epsilon] \quad [\mathbb{P}]\text{-}a.s. \end{split}$$

Hence,

$$\begin{split} &|\mathbb{E}_{\mathbb{P}_{\chi_{n}}^{*}}h([\xi_{n}^{*},\zeta_{n}^{*}]) - \mathbb{E}_{\mathbb{P}}h([\xi_{0},\zeta_{0}])| \\ &\leq |\mathbb{E}_{\mathbb{P}_{\chi_{n}}^{*}}h([\xi_{n}^{*},\zeta_{n}^{*}]) - \mathbb{E}_{\mathbb{P}_{\chi_{n}}^{*}}h([\xi_{n}^{*},\zeta_{0}])| + |\mathbb{E}_{\mathbb{P}_{\chi_{n}}^{*}}h([\xi_{n}^{*},\zeta_{0}]) - \mathbb{E}_{\mathbb{P}}h([\xi_{0},\zeta_{0}])| \\ &\leq K\epsilon + 2M\mathbb{P}_{\chi_{n}}^{*}[\|[\xi_{n}^{*},\zeta_{n}^{*}] - [\xi_{n}^{*},\zeta_{0}]\| \geq \epsilon] \\ &+ |\mathbb{E}_{\mathbb{P}_{\chi_{n}}^{*}}h([\xi_{n}^{*},\zeta_{0}]) - \mathbb{E}_{\mathbb{P}}h([\xi_{0},\zeta_{0}])| \quad [\mathbb{P}]\text{-}a.s. \end{split}$$
(A4)

We take the limit in this expression as $n \to \infty$. Since (A4) holds for arbitrary $\epsilon > 0$, the second term will go to zero $[\mathbb{P}]$ -almost surely due to (A3) and Definition 3. The third term (does not depend on ϵ) will also converge to zero $[\mathbb{P}]$ -almost surely by (A2). Thus,

$$\lim_{n\to\infty} |\mathbb{E}_{\mathbb{P}^*_{\chi_n}} h([\boldsymbol{\xi}^*_n,\boldsymbol{\zeta}^*_n]) - \mathbb{E}_{\mathbb{P}} h([\boldsymbol{\xi}_0,\boldsymbol{\zeta}_0])| \le K\epsilon \quad [\mathbb{P}]\text{-}a.s.$$

Since ϵ was arbitrary, we conclude that the limit must in fact be equal to zero $[\mathbb{P}]$ -almost surely. Therefore, (*i*) is proved.

In order to prove result (*viii*), consider again $\|[\xi_n^*, \zeta_n^*] - [\xi_n^*, \zeta_0]\| = \|\zeta_n^* - \zeta_0\|$. This expression converges in probability $\mathbb{P}_{\chi_n}^*$ in probability \mathbb{P} to zero due to assumption (6). Thus

$$[\boldsymbol{\xi}_n^*, \boldsymbol{\zeta}_n^*] - [\boldsymbol{\xi}_n^*, \boldsymbol{\zeta}_0] \xrightarrow[n \to \infty]{\mathbb{P}^*_{\boldsymbol{\chi}_n}(\mathbb{P})} \boldsymbol{0}, \tag{A5}$$

Which can be rewritten according to Definition 3 as

$$\forall \epsilon > 0: \left\{ \mathbb{P}^*_{\boldsymbol{\chi}_n}[\|[\boldsymbol{\xi}^*_n, \boldsymbol{\zeta}^*_n] - [\boldsymbol{\xi}^*_n, \boldsymbol{\zeta}_0]\| \ge \epsilon] \xrightarrow[n \to \infty]{\mathbb{P}} 0 \right\}$$

Let us fix $\epsilon > 0$. Similarly as in the proof of part (*i*), it can be demonstrated that

$$\mathbb{E}_{\mathbb{P}^*_{\boldsymbol{\chi}n}} f([\boldsymbol{\xi}_n^*, \boldsymbol{\zeta}_0]) - \mathbb{E}_{\mathbb{P}} f([\boldsymbol{\xi}_0, \boldsymbol{\zeta}_0]) \xrightarrow[n \to \infty]{\mathbb{P}} 0,$$
(A6)

When convergence $[\mathbb{P}]$ -almost surely is just replaced by convergence in probability \mathbb{P} . Indeed, using inequality (A4), we get for arbitrary $\tau > 0$ (and above chosen fixed and sufficiently small $\epsilon > 0$)

$$\begin{split} & \mathbb{P}\Big[|\mathbb{E}_{\mathbb{P}^*_{\chi_n}}h([\boldsymbol{\xi}^*_n,\boldsymbol{\zeta}^*_n]) - \mathbb{E}_{\mathbb{P}}h([\boldsymbol{\xi}_0,\boldsymbol{\zeta}_0])| \geq \tau\Big] \\ & \leq \mathbb{P}[K\epsilon \geq \tau] + \mathbb{P}\Big[2M\mathbb{P}^*_{\chi_n}[\|[\boldsymbol{\xi}^*_n,\boldsymbol{\zeta}^*_n] - [\boldsymbol{\xi}^*_n,\boldsymbol{\zeta}_0]\| \geq \epsilon] \geq \tau\Big] \\ & + \mathbb{P}\Big[|\mathbb{E}_{\mathbb{P}^*_{\chi_n}}h([\boldsymbol{\xi}^*_n,\boldsymbol{\zeta}_0]) - \mathbb{E}_{\mathbb{P}}h([\boldsymbol{\xi}_0,\boldsymbol{\zeta}_0])| \geq \tau\Big]. \end{split}$$

If we take the limit in previous inequality as n goes to infinity, the first term is zero and the second term will go to zero due to (A5) and Definition 3. The third term will also converge to zero by (A6). Thus

$$\lim_{n\to\infty}\mathbb{P}\Big[\big|\mathbb{E}_{\mathbb{P}^*_{\chi_n}}h([\boldsymbol{\xi}^*_n,\boldsymbol{\zeta}^*_n])-\mathbb{E}_{\mathbb{P}}h([\boldsymbol{\xi}_0,\boldsymbol{\zeta}_0])\big|\geq\tau\Big]=0.$$

Since τ was arbitrary, (*viii*) is proved.

Assertions (*ii*)–(*vii*) are just corollaries of (*i*), when continuous mapping theorem (see, e.g., Theorems 2 and 3 [55]) is applied. Similarly, assertions (*ix*)–(*xiv*) are consequences of (*viii*). \Box

Proof of Lemma 2. See [56] (Theorem 1). \Box

Proof of Lemma 3. See [57] (Theorem 4.1). \Box

Proof of Lemma 4. See [23] (Theorem 5.2). \Box

Proof of Lemma 5. See [58] or [26] (Corollary 3.2.1). □

Proof of Corollary 1. Since a functional distributional limit (a convergence in Skorokhod space) implies the pointwise distributional limit, this corollary is just a special case of Lemma 5, when $W_n(1)$ is considered. \Box

Proof of Lemma 6. See [59] (Corollary 4). \Box

Proof of Corollary 2. We show that Assumptions (20) and (21) implies the Lindeberg Condition (19) from Lemma 6.

The first step is to show that Conditions (20) and (21) implies the so-called Lyapunov condition, i.e., having fixed $\omega > 0$:

$$\frac{1}{\zeta_n^{2+\omega}}\sum_{i=1}^n \mathbb{E}|\xi_i|^{2+\omega} \leq \frac{1}{\zeta_n^{2+\omega}}\sum_{i=1}^n \sup_{\iota \in \mathbb{N}} \mathbb{E}|\xi_\iota|^{2+\omega} = \frac{n}{\zeta_n^{2+\omega}}\sup_{\iota \in \mathbb{N}} \mathbb{E}|\xi_\iota|^{2+\omega} \to 0, \quad n \to \infty$$

Now, the Lyapunov condition $\lim_{n\to\infty} \zeta_n^{-2-\omega} \sum_{i=1}^n \mathbb{E} |\xi_i|^{2+\omega} = 0$ holds and we fix $\delta > 0$. Since $|\xi_i| > \delta \zeta_n$ implies $|\xi_i/\delta \zeta_n|^{\omega} > 1$, we obtain

$$\begin{split} \frac{1}{\varsigma_n^2} \sum_{i=1}^n \mathbb{E} \xi_i^2 \mathcal{I}\{|\xi_i| > \delta \varsigma_n\} &\leq \frac{1}{\delta^{\omega} \varsigma_n^{2+\omega}} \sum_{i=1}^n \mathbb{E} |\xi_i|^{2+\omega} \mathcal{I}\{|\xi_i| > \delta \varsigma_n\} \\ &\leq \frac{1}{\delta^{\omega} \varsigma_n^{2+\omega}} \sum_{i=1}^n \mathbb{E} |\xi_i|^{2+\omega} \to 0, \quad n \to \infty. \end{split}$$

Proof of Theorem 2. The SLLN applied for independent random variables together with Assumption (23) lead to

$$n^{-1}\sum_{i=1}^{n} (\xi_i - \mathbb{E}_{\mathbb{P}}\xi_i) \xrightarrow[n \to \infty]{[\mathbb{P}]\text{-}a.s.} 0.$$

Markov's inequality with (23) implies uniform equiboundedness in probability \mathbb{P} of ξ_n^2 . The conditional variance of the bootstrapped sample mean goes to zero as *n* increases to infinity, because

$$\mathbb{V}\mathrm{ar}_{\mathbb{P}_{\xi}^{*}}\left(n^{-1}\sum_{i=1}^{n}\xi_{i}^{*}\right) = n^{-1}\mathbb{V}\mathrm{ar}_{\mathbb{P}_{\xi}^{*}}\xi_{1}^{*} = n^{-1}\left[\mathbb{E}_{\mathbb{P}_{\xi}^{*}}\xi_{1}^{*2} - (\mathbb{E}_{\mathbb{P}_{\xi}^{*}}\xi_{1}^{*})^{2}\right]$$
$$= n^{-1}\left[\sum_{k=1}^{n}n^{-1}\xi_{k}^{2} - \left(\sum_{k=1}^{n}n^{-1}\xi_{k}^{k}\right)^{2}\right] = \mathcal{O}_{\mathbb{P}}(n^{-1}), \quad n \to \infty.$$

Hence, the weak law of large numbers in the "starred" world (for resampled variables) provides

$$n^{-1}\sum_{i=1}^{n} \left(\xi_{i}^{*} - \mathbb{E}_{\mathbb{P}_{\xi}^{*}}\xi_{i}\right) = n^{-1}\sum_{i=1}^{n}\xi_{i}^{*} - n^{-1}\sum_{i=1}^{n}\xi_{i} \xrightarrow{\mathbb{P}_{\xi}^{*}(\mathbb{P})}{n \to \infty} 0,$$

Because ξ_i^* are conditionally IID. \Box

Proof of Theorem 3. The sequence $\{\xi_n\}_{n=1}^{\infty}$ is uniformly bounded in probability \mathbb{P} , because (25) is assumed. By the assumptions for $\{\xi_n\}_{n=1}^{\infty}$ and Lemma 2, it is known that a SLLN holds for the sample mean $n^{-1}\sum_{i=1}^{n} \xi_i$. By Corollary A2 from [12], it follows that

$$\operatorname{Var}_{\mathbb{P}^*_{\xi}}\left(n^{-1}\sum_{i=1}^n\xi^*_i\right) = \mathcal{O}_{\mathbb{P}}(n^{-1}), \quad n \to \infty.$$

Thus the claim holds due to Lemma A1 by [12], where

$$\mathbb{E}_{\mathbb{P}^*_{\xi}} n^{-1} \sum_{i=1}^n \xi^*_i = n^{-1} \sum_{i=1}^n \xi_i + o_{\mathbb{P}}(n^{-1/2}), \quad n \to \infty.$$

Proof of Theorem 4. The proof is the same as the proof of previous Lemma 3 except one detail—the SLLN for φ -mixing (Lemma 3) has to be applied instead of Lemma 2 for α -mixing.

Proof of Theorem 5. See [60]. \Box

Proof of Theorem 6. The Lyapunov condition for sequence of random variables $\{\xi_n\}_{n=1}^{\infty}$ is satisfied due to (28) and (29), i.e., for fixed $\omega > 0$:

$$\frac{1}{\varsigma_n^{2+\omega}}\sum_{i=1}^n \mathbb{E}|\xi_i|^{2+\omega} \leq \frac{1}{\varsigma_n^{2+\omega}}\sum_{i=1}^n \sup_{\iota \in \mathbb{N}} \mathbb{E}|\xi_\iota|^{2+\omega} = \frac{n}{\varsigma_n^{2+\omega}}\sup_{\iota \in \mathbb{N}} \mathbb{E}|\xi_\iota|^{2+\omega} \to 0, \quad n \to \infty.$$

Thereupon, the CLT for $\{\xi_n\}_{n=1}^{\infty}$ holds and

$$\sup_{x \in \mathbb{R}} \left| \mathbb{P}\left[\frac{n}{\sqrt{\zeta_n^2}} \bar{\zeta}_n \le x \right] - \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} \exp\left\{ -\frac{t^2}{2} \right\} dt \right| \xrightarrow[n \to \infty]{} 0$$

Henceforth, to prove this theorem, it suffices to show the following three statements:

$$\sup_{x\in\mathbb{R}}\left|\mathbb{P}_{\xi}^{*}\left[\frac{\sqrt{n}}{\sqrt{\mathbb{V}\mathrm{ar}}_{\mathbb{P}_{\xi}^{*}}\tilde{\xi}_{1}^{*}}\left(\bar{\xi}_{n}^{*}-\mathbb{E}_{\mathbb{P}_{\xi}^{*}}\bar{\xi}_{n}^{*}\right)\leq x\right]-\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{t^{2}}{2}\right\}\mathrm{d}t\left|\frac{\mathbb{P}}{n\to\infty}0;\right.$$

(ii)

(i)

$$\operatorname{\mathbb{V}ar}_{\mathbb{P}^*_{\xi}} \xi_1^* - n^{-1} \zeta_n^2 \xrightarrow[n \to \infty]{[\mathbb{P}]\text{-a.s.}} 0;$$

(iii)

$$\mathbb{E}_{\mathbb{P}^*_{\tau}}\bar{\xi}^*_n = \bar{\xi}_n \quad [\mathbb{P}]\text{-}a.s.$$

Proving (iii) is trivial, because the bootstrapped variables $\{\xi_n^*\}_{n=1}^{\infty}$ are conditionally IID and, therefore,

$$\mathbb{E}_{\mathbb{P}^*_{\xi}}\bar{\xi}^*_n = \mathbb{E}_{\mathbb{P}^*_{\xi}}\xi^*_1 = n^{-1}\sum_{i=1}^n \xi_i = \bar{\xi}_n \quad [\mathbb{P}]\text{-}a.s$$

Let us calculate the conditional variance of the bootstrapped ξ_1^* :

$$\operatorname{Var}_{\mathbb{P}^*_{\xi}} \xi_1^* = \mathbb{E}_{\mathbb{P}^*_{\xi}} \xi_1^{*2} - (\mathbb{E}_{\mathbb{P}^*_{\xi}} \xi_1^*)^2 = n^{-1} \sum_{i=1}^n \xi_i^2 - \left(n^{-1} \sum_{i=1}^n \xi_i\right)^2 \quad [\mathbb{P}]\text{-}a.s.$$

The strong law of large numbers for independent non-identically distributed random variables with (28) provide

$$\bar{\xi}_n - n^{-1} \sum_{i=i}^n \mathbb{E}_{\mathbb{P}} \xi_i = \bar{\xi}_n \xrightarrow[n \to \infty]{[\mathbb{P}]^{-a.s.}} 0$$

and

$$0 \xleftarrow{[\mathbb{P}]^{-a.s.}}_{n \to \infty} n^{-1} \sum_{i=1}^{n} \xi_{i}^{2} - \left(n^{-1} \sum_{i=1}^{n} \xi_{i} \right)^{2} - n^{-1} \sum_{i=1}^{n} \mathbb{E}_{\mathbb{P}} \xi_{i}^{2} = \mathbb{V} \operatorname{ar}_{\mathbb{P}_{\xi}^{*}} \xi_{1}^{*} - n^{-1} \zeta_{n}^{2}.$$

The last result of the SLLN is true, because (28) implies

$$\sum_{n=1}^{\infty} \frac{\mathbb{V}\mathrm{ar}_{\mathbb{P}} \xi_n^2}{n^2} \leq \sum_{n=1}^{\infty} \frac{\mathbb{E}_{\mathbb{P}} \xi_n^4}{n^2} \leq \left[\sup_{\iota \in \mathbb{N}} \mathbb{E}_{\mathbb{P}} \xi_\iota^4 \right] \sum_{n=1}^{\infty} n^{-2} < \infty.$$

Thus (*ii*) is proved.

The Berry–Esseen-Katz Theorem 5 with $g(x) = |x|^{\epsilon}$, $\epsilon > 0$ for the bootstrapped sequence of IID (with respect to \mathbb{P}^*) random variables $\{\xi_n^*\}_{n=1}^{\infty}$ results into

$$\begin{split} \sup_{x \in \mathbb{R}} \left| \mathbb{P}_{\xi}^{*} \left[\frac{\sqrt{n}}{\sqrt{\mathbb{V}\mathrm{ar}}_{\mathbb{P}_{\xi}^{*}} \tilde{\zeta}_{1}^{*}} \left(\tilde{\zeta}_{n}^{*} - \mathbb{E}_{\mathbb{P}_{\xi}^{*}} \tilde{\zeta}_{n}^{*} \right) \leq x \right] - \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp\left\{ -\frac{t^{2}}{2} \right\} \mathrm{d}t \right| \\ \leq Cn^{-\epsilon/2} \mathbb{E}_{\mathbb{P}_{\xi}^{*}} \left| \frac{\tilde{\zeta}_{1}^{*} - \mathbb{E}_{\mathbb{P}_{\xi}^{*}} \tilde{\zeta}_{1}^{*}}{\sqrt{\mathbb{V}\mathrm{ar}}_{\mathbb{P}_{\xi}^{*}} \tilde{\zeta}_{1}^{*}} \right|^{2+\epsilon} \quad [\mathbb{P}]\text{-}a.s., \quad (A7) \end{split}$$

where C > 0 is an absolute constant.

The Minkowski inequality and Jensen's inequality provide an upper bound for the nominator from the right-hand side of (A7):

$$\begin{split} & \mathbb{E}_{\mathbb{P}_{\xi}^{*}} |\xi_{1}^{*} - \mathbb{E}_{\mathbb{P}_{\xi}^{*}} \xi_{1}^{*}|^{2+\epsilon} = n^{-1} \sum_{i=1}^{n} \left| \xi_{i} - n^{-1} \sum_{j=1}^{n} \xi_{j} \right|^{2+\epsilon} \\ & \leq n^{-1} \left\{ \left(\sum_{i=1}^{n} |\xi_{i}|^{2+\epsilon} \right)^{1/(2+\epsilon)} + n^{-(1+\epsilon)/(2+\epsilon)} \left| \sum_{j=1}^{n} \xi_{j} \right| \right\}^{2+\epsilon} \\ & \leq 2^{1+\epsilon} n^{-1} \sum_{i=1}^{n} |\xi_{i}|^{2+\epsilon} + 2^{1+\epsilon} \left| n^{-1} \sum_{i=1}^{n} \xi_{i} \right|^{2+\epsilon} \quad [\mathbb{P}]\text{-}a.s. \end{split}$$

The right-hand side of the previously derived upper bound is uniformly bounded in probability \mathbb{P} , because of Markov's inequality and (28). In very deed, for fixed $\tau > 0$

$$\mathbb{P}\left[n^{-1}\sum_{i=1}^{n}|\xi_{i}|^{2+\epsilon} \geq \tau\right] \leq \tau^{-1}n^{-1}\sum_{i=1}^{n}\mathbb{E}_{\mathbb{P}}|\xi_{i}|^{2+\epsilon} \leq \tau^{-1}\sup_{\iota\in\mathbb{N}}\mathbb{E}_{\mathbb{P}}|\xi_{\iota}|^{2+\epsilon} < \infty, \quad \forall n\in\mathbb{N}$$

and

$$\mathbb{P}\left[\left|n^{-1}\sum_{i=1}^{n}\xi_{i}\right| \geq \tau\right] \leq \tau^{-1}n^{-1}\mathbb{E}_{\mathbb{P}}\left|\sum_{i=1}^{n}\xi_{i}\right| \leq \tau^{-1}\sup_{\iota\in\mathbb{N}}\mathbb{E}_{\mathbb{P}}|\xi_{\iota}| < \infty, \quad \forall n \in \mathbb{N}.$$

Since $\mathbb{E}_{\mathbb{P}^*_{\xi}} |\xi_1^* - \mathbb{E}_{\mathbb{P}^*} \xi_1^*|^{2+\epsilon}$ is bounded in probability \mathbb{P} uniformly over *n* and the denominator of the right-hand side of (A7) is uniformly bounded away from zero due to (29), then the left-hand side of (A7) converges in probability \mathbb{P} to zero as *n* tends to infinity. So, (*i*) is proved as well. \Box

Proof of Theorem 7. According to Cramér–Wold theorem, it is sufficient to ensure that all the assumptions of one-dimensional bootstrap CLT 6 are valid for any linear combination of the elements of random vector ξ_n , $n \in \mathbb{N}$.

For arbitrary fixed $\mathbf{t} \in \mathbb{R}^q$ using Jensen's inequality, we get

$$\sup_{n\in\mathbb{N}}\mathbb{E}_{\mathbb{P}}|\mathbf{t}^{\top}\boldsymbol{\xi}_{n}|^{4}\leq q^{3}\sup_{n\in\mathbb{N}}\sum_{j=1}^{q}t_{j}^{4}\mathbb{E}_{\mathbb{P}}|\boldsymbol{\xi}_{j,n}|^{4}\leq q^{4}\max_{j=1,...,q}t_{j}^{4}\sup_{n\in\mathbb{N}}\mathbb{E}_{\mathbb{P}}|\boldsymbol{\xi}_{j,n}|^{4}<\infty.$$

Hence, Assumption (33) implies Assumption (28) for random variables $\mathbf{t}^{\top} \boldsymbol{\xi}_{n}$, $n \in \mathbb{N}$.

Similar situation arises, when assumption (34) implies assumption (29) for such an arbitrary linear combination, i.e., positive definiteness of matrix Γ yields

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\mathbb{V}\mathrm{ar}_{\mathbb{P}}\mathbf{t}^{\top}\boldsymbol{\xi}_{i}=\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\mathbf{t}^{\top}(\mathbb{V}\mathrm{ar}_{\mathbb{P}}\boldsymbol{\xi}_{i})\mathbf{t}=\mathbf{t}^{\top}\left(\lim_{n\to\infty}\frac{1}{n}\boldsymbol{\Gamma}_{n}\right)\mathbf{t}=\mathbf{t}^{\top}\boldsymbol{\Gamma}\mathbf{t}>0.$$

Finally, we need to realize that (34) holds, (35) has already been proved above, $\{\xi_i^*\}_{i=1}^n$ are conditionally IID, and

$$\mathbb{E}_{\mathbb{P}_{\Xi}^*}\boldsymbol{\xi}_1^* = n^{-1}\sum_{i=1}^n \boldsymbol{\xi}_i = \bar{\boldsymbol{\xi}}_n.$$

Proof of Theorem 8. The proof of this theorem goes along the lines the proof of Theorem 6 except three facts: the bootstrap WLLN for α -mixing–Lemma 3 is used; the CLT for α -mixing–Corollary 1 is employed; and another covariance inequality is applied, i.e., Lemma 1.2.5 by [26]. \Box

Proof of Theorem 9. The proof of this theorem remains almost the same as the proof of Theorem 8 with the only exception that the following tools have to be used instead: the CLT for φ -mixing observations, i.e., Corollary 2; a different covariance inequality, i.e., Lemma 1.2.8 by [26]; and a suitable bootstrap WLLN, i.e., Lemma 4.

References

- 1. Efron, B. Bootstrap methods: Another look at the jackknife. Ann. Stat. 1979, 7, 1–26.
- 2. Bickel, P.J.; Freedman, D.A. Some asymptotic theory for the bootstrap. Ann. Stat. 1981, 9, 1196–1217. [CrossRef]
- 3. Efron, B.; Tibshirani, R. An Introduction to the Bootstrap; Chapman and Hall: London, UK, 1990.
- 4. Efron, B.; Hastie, T. Computer Age Statistical Inference, Student ed.; Cambridge University Press: Cambridge, UK, 2013.
- 5. Hall, P. The Bootstrap and Edgeworth Expansion; Springer: New York, NY, USA, 2013.
- 6. Lahiri, S.N. Resampling Methods for Dependent Data; Springer: New York, NY, USA, 2003.
- 7. Politis, D.N.; Romano, J.P.; Wolf, M. Subsampling; Springer: New York, NY, USA, 1999.
- 8. Künsch, H.R. The jacknife and the bootstrap for general stationary observations. Ann. Stat. 1989, 17, 1217–1241. [CrossRef]
- 9. Liu, R.Y.; Singh, K. Moving blocks jackknife and bootstrap capture weak dependence. In *Proceedings of the Exploring the Limits of Bootstrap*; Lapage, R., Billard, L., Eds.; Wiley: New York, NY, USA, 1992; pp. 225–248.
- 10. Lahiri, S.N. Non-strong mixing autoregressive processes. *Stat. Probabil. Lett.* **1991**, *11*, 335–341.
- 11. Politis, D.N.; Romano, J.P. A general resampling scheme for triangular arrays of *α*-mixing random variables with application to the problem of spectral density estimation. *Ann. Stat.* **1992**, *20*, 1985–2007. [CrossRef]
- 12. Fitzenberger, B. The moving block bootstrap and robust inference for linear least squares and quantile regression. *J. Econometr.* **1998**, *82*, 235–287. [CrossRef]
- 13. Maciak, M.; Peštová, B.; Pešta, M. Structural breaks in dependent, heteroscedastic, and extremal panel data. *Kybernetika* 2019, 54, 1106–1121. [CrossRef]
- 14. Pesarin, F.; Salmaso, L. Permutation Tests for Complex Data: Theory, Applications and Software; Wiley: New York, NY, USA, 2010.
- 15. Pešta, M. Total least squares and bootstrapping with application in calibration. *Statistics* **2013**, 47, 966–991. [CrossRef]
- 16. Pešta, M. Block bootstrap for dependent errors-in-variables. Commun. Stat. A-Theory 2017, 46, 1871–1897. [CrossRef]
- 17. Hall, P.; Horowitz, J.L.; Jing, B.Y. On blocking rules for the bootstrap with dependent data. Biometrika 1995, 82, 561–574.
- 18. Politis, D.N.; White, H. Automatic block-length selection for the dependent bootstrap. Econometr. Rev. 2004, 23, 53–70. [CrossRef]
- 19. Lahiri, S.; Furukawa, K.; Lee, Y.D. A nonparametric plug-in rule for selecting optimal block lengths for block bootstrap methods. *Stat. Methodol.* **2007**, *4*, 292–321. [CrossRef]
- 20. Kirch, C. Resampling Methods for the Change Analysis of Dependent Data. Ph.D. Thesis, University of Cologne, Cologne, Germany, 2006.
- 21. Peštová, B.; Pešta, M. Abrupt change in mean using block bootstrap and avoiding variance estimation. *Comput. Stat.* **2018**, *33*, 413–441.
- 22. Billingsley, P. Convergence of Probability Measures, 2nd ed.; John Wiley & Sons: New York, NY, USA, 1999.
- 23. Bradley, R.C. Basic properties of strong mixing conditions. A survey and some open questions. Probab. Surv. 2005, 2, 107–144.
- 24. Gijbels, I.; Omelka, M.; Pešta, M.; Veraverbeke, N. Score tests for covariate effects in conditional copulas. *J. Multivar. Anal.* 2017, 159, 111–133.
- 25. Rosenblatt, M. A central limit theorem and a strong mixing condition. Proc. Natl. Acad. Sci. USA 1956, 42, 43–47. [CrossRef]
- 26. Lin, Z.; Lu, C. Limit Theory for Mixing Dependent Random Variables; Springer: New York, NY, USA, 1997.
- 27. Ibragimov, I.A. Some limit theorems for stochastic processes stationary in the strict sense. *Dokl. Akad. Nauk. SSSR* **1959**, *125*, 711–714. (In Russian)
- 28. Pešta, M. Asymptotics for weakly dependent errors-in-variables. *Kybernetika* 2013, 49, 692–704.
- 29. Anderson, T.W. An Introduction to Multivariate Statistical Analysis; John Wiley & Sons: New York, NY, USA, 1958.
- 30. Billingsley, P. Convergence of Probability Measures, 1st ed.; John Wiley & Sons: New York, NY, USA, 1968.
- 31. Doob, J.L. Stochastic Processes; John Wiley & Sons: New York, NY, USA, 1953.
- 32. Ibragimov, I.A.; Linnik, Y.V. Independent and Stationary Sequences of Random Variables; Wolters-Noordhoff: Amsterdam, The Netherlands, 1971.
- 33. Rosenblatt, M. Markov Processes: Structure and Asymptotic Behavior; Springer: Berlin/Heidelberg, Germany, 1971.
- Antoch, J.; Hušková, M.; Prášková, Z. Effect of dependence on statistics for determination of change. J. Stat. Plan. Infer. 1997, 60, 291–310. [CrossRef]
- 35. Andrews, D.W.K. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica* **1991**, *59*, 817–858. [CrossRef]
- 36. Belyaev, Y.K. *Bootstrap, Resampling and Mallows Metric. Lecture Notes 1*; Institute of Mathematical Statistics, UmeåUniversity: Umeå, Sweden, 1995.
- 37. Singh, K. On the asymptotic accuracy of Efron's bootstrap. Ann. Stat. 1981, 9, 1187–1195. [CrossRef]
- 38. Peštová, B.; Pešta, M. Testing structural changes in panel data with small fixed panel size and bootstrap. *Metrika* **2015**, *78*, 665–689.

- 39. Peštová, B.; Pešta, M. Erratum to: Testing structural changes in panel data with small fixed panel size and bootstrap. *Metrika* **2016**, *79*, 237–238. [CrossRef]
- 40. Peštová, B.; Pešta, M. Change point estimation in panel data without boundary issue. Risks 2017, 5, 7. [CrossRef]
- 41. Culpepper, S.A.; Balamuta, J.J. Inferring latent structure in polytomous data with a higher-order diagnostic model. *Multivar. Behav. Res.* **2021**, 1–19. [CrossRef]
- 42. Maciak, M.; Mizera, I.; Pešta, M. Functional profile techniques for claims reserving. ASTIN Bull. 2022, 52, 449–482.
- 43. Gerthofer, M.; Pešta, M. Stochastic claims reserving in insurance using random effects. *Prague Econ. Pap.* **2017**, *26*, 542–560. [CrossRef]
- 44. McNeil, A.J. Estimating the tails of loss severity distributions using extreme value theory. *ASTIN Bull.* **1997**, 27, 117–137. [CrossRef]
- 45. Pešta, M. Changepoint in error-prone relations. *Mathematics* **2021**, *9*, 89. [CrossRef]
- 46. Pešta, M.; Wendler, M. Nuisance-parameter-free changepoint detection in non-stationary series. Test 2020, 29, 379–408.
- Maciak, M.; Pešta, M.; Peštová, B. Changepoint in dependent and non-stationary panels. *Stat. Pap.* 2020, *61*, 1385–1407. [CrossRef]
 Pešta, M.; Peštová, B.; Maciak, M. Changepoint estimation for dependent and non-stationary panels. *Appl. Math.-Czech.* 2020, *65*,
- 299–310. [CrossRef]
 49. Pešta, M.; Hudecová, Š. Asymptotic consistency and inconsistency of the chain ladder. *Insur. Math. Econ.* 2012, *51*, 472–479. [CrossRef]
- 50. Hudecová, Š.; Pešta, M. Modeling dependencies in claims reserving with GEE. Insur. Math. Econ. 2013, 53, 786–794.
- Pešta, M.; Okhrin, O. Conditional least squares and copulae in claims reserving for a single line of business. *Insur. Math. Econ.* 2014, 56, 28–37. [CrossRef]
- 52. Maciak, M.; Okhrin, O.; Pešta, M. Infinitely stochastic micro reserving. Insur. Math. Econ. 2021, 100, 30-58.
- 53. Belyaev, Y.K.; Sjöstedt-de Luna, S. Weakly approaching sequences of random distributions. *J. Appl. Probab.* 2000, 37, 807–822. [CrossRef]
- 54. Zagdanski, A. On the construction and properties of bootstrap-t prediction intervals for stationary time series. *Probab. Math. Statist.* **2005**, 25, 133–153.
- 55. van der Vaart, A.W. Asymptotic Statistics; Cambridge University Press: New York, NY, USA, 1998.
- 56. Chen, X.; Wu, Y. Strong law for mixing sequence. Acta Math. Appl. Sin. 1989, 5, 367–371.
- Xuejun, W.; Shuhe, H.; Yan, S.; Wenzhi, Y. Moment inequalities for *φ*-mixing sequences and its applications. *J. Inequal. Appl.* 2009, 2009, 12. [CrossRef]
- Herrndorf, N. A functional central limit theorem for strongly mixing sequence of random variables. *Probab. Theory Relat. Fields* 1985, 69, 541–550. [CrossRef]
- 59. Utev, S.A. The central limit theorem for *φ*-mixing arrays of random variables. *Theory Probab. Appl.* **1990**, *35*, 131–139. [CrossRef]
- 60. Katz, M.L. Note on the Berry–Esseen theorem. Ann. Math. Stat. 1963, 34, 1107–1108.