Bootstrap for functions of associated random variables with applications

MANSI GARG AND ISHA DEWAN

Indian Statistical Institute, Delhi Centre
7, SJSS Marg, New Delhi–110 016, India
Bootstrap for functions of associated random variables with applications

Mansi Garg and Isha Dewan
Indian Statistical Institute
New Delhi-110016 (India)
mansibirla@gmail.com and ishadewan@gmail.com

Abstract

Let \( \{X_n, n \geq 1\} \) be a sequence of stationary associated random variables. In this paper, we obtain consistent estimators of the distribution function and the variance of the sample mean based on \( \{g(X_n), n \geq 1\} \), \( g: \mathbb{R} \rightarrow \mathbb{R} \) using Circular Block Bootstrap (CBB). We extend these results to derive consistent estimators of the distribution function and the variance of U-statistics. As applications, we obtain interval estimators for L-moments. We also discuss consistent point estimators for L-moments. Finally, as an illustration, we obtain point estimators and confidence intervals for L-moments of a stationary autoregressive process with a minification structure which is fitted to a hydrological dataset.

Keywords: Associated random variables; Circular Block Bootstrap; L-moments; Hardy-Krause variation; U-statistics.

1 Introduction

Following the introduction of bootstrap by Efron (1979), several variants of bootstrap procedure and their extensions to dependent data have been widely discussed in literature. Initial publications on block resampling techniques include Künsch(1989), Lui and Singh (1992), Politis and Romano (1992, 1994), among others. Some recent publications on bootstrap and subsampling techniques for dependent data are by Shao (2010), Doukhan et. al. (2011), Hwang and Shin (2012), and Doukhan et. al. (2015). A compilation of results on bootstrap and other resampling procedures can be found in Efron and Tibshirani (1994), Lahiri (2003), and Shao and Tu (2012). These nonparametric techniques are often used for constructing tests of hypothesis and confidence intervals of parameters.

In recent years, results on bootstrap for U-statistics based on dependent random variables have also been presented by several authors. Dehling and Wendler (2010), and Sharipov and

In this paper, we discuss consistency of the Circular Block Bootstrap technique and estimation of L-moments for \( \{X_n, n \geq 1\} \), where \( \{X_n, n \geq 1\} \) is a sequence of stationary associated random variables. Apropos our discussion, we give the following.

**Definition 1.1.** A finite collection of random variables \( \{X_j, 1 \leq j \leq n\} \) is said to be associated, if for any choice of component-wise nondecreasing functions \( h_1, h_2 : \mathbb{R}^n \rightarrow \mathbb{R} \), we have,

\[
\text{Cov}(h_1(X_1, \ldots, X_n), h_2(X_1, \ldots, X_n)) \geq 0
\]

whenever it exists. An infinite collection of random variables \( \{X_n, n \geq 1\} \) is associated if every finite sub-collection is associated.

Introduced by Esary et al. (1967), associated random variables have been widely used in reliability studies, statistical mechanics, and percolation theory. Any set of independent random variables is associated. Nondecreasing functions of associated random variables are associated, for example, order statistics corresponding to a finite set of independent random variables, the moving average process \( \{X_n, n \geq 1\} \), where \( X_n = a_0 \epsilon_n + \ldots + a_q \epsilon_{n-q} \), \( \epsilon_n \) are independent random variables and \( a_0, \ldots, a_q \) have the same sign. A detailed presentation of theory and results for associated random variables can be found in Bulinski and Shashkin (2007, 2009), Prakasa Rao (2012), and Oliveira (2012).

This paper is organized as follows. In Section 2, we give results and definitions that will be required to prove our main results in Sections 3 and 4. In Section 3, we obtain consistent estimators of the variance and the distribution function of the sample mean based on \( \{g(X_n), n \geq 1\} \) \( g : \mathbb{R} \rightarrow \mathbb{R} \) using Circular Block Bootstrap (CBB). In Section 4, we extend the results of Section 3 to obtain consistent estimators of variance and distribution function of U-statistics based on \( \{X_n, n \geq 1\} \). In Section 5, we apply the results discussed in Sections 2 – 4 to obtain point estimators and confidence intervals for L-moments of \( \{X_n, n \geq 1\} \). As an example, we obtain point estimators and confidence intervals of the first three L-moments of a minification process. Simulation results are also presented along. The peak flow data (1883 – 2014) of the Thames river, Kingston, UK is presented as a case study. In Section 6, we give a brief summary of the results obtained and discuss our intended future work.
2 Preliminaries

In this section, we give results and definitions which will be needed to prove our main results given in Sections 3 and 4.

**Definition 2.1.** (Newman (1984)) If \( g \) and \( \tilde{g} \) are two real-valued functions on \( \mathbb{R}^k \), \( k \in \mathbb{N} \), then \( g \ll \tilde{g} \) iff \( \tilde{g} + g \) and \( \tilde{g} - g \) are both coordinate-wise nondecreasing. If \( g \ll \tilde{g} \), then \( \tilde{g} \) will be coordinate-wise nondecreasing.

Define, \( Y_n = g(X_n) \), and \( \tilde{Y}_n = \tilde{g}(X_n) \), \( g \ll \tilde{g}, \ n \geq 1 \).

**Lemma 2.2.** (Newman (1984)) Define,

\[
\sigma^2 = \text{Var}(Y_1) + 2 \sum_{j=2}^{\infty} \text{Cov}(Y_1, Y_j).
\]

Let \( \sigma^2 > 0 \) and \( \sum_{j=1}^{\infty} \text{Cov}(\tilde{Y}_1, \tilde{Y}_j) < \infty \). Then,

\[
\frac{1}{\sqrt{n}\sigma} \sum_{j=1}^{n} (Y_j - E(Y_j)) \overset{\text{d}}{\rightarrow} N(0, 1) \quad \text{as} \quad n \to \infty.
\]

We next give a Central limit theorem for U-statistics based on \( \{X_n, n \geq 1\} \). We assume that the underlying kernels of the U-statistics are functions of bounded Hardy-Krause variation. This result was obtained in Garg and Dewan (2015).

Before proceeding, we give the following definitions discussing the concept of Hardy-Krause variation (For further discussions on this topic, see Beare (2009) and the references therein.)

**Definition 2.3.** The Vitali variation of a function \( f : [a, b] \to \mathbb{R} \), is defined as \( ||f||_V = \sup \sum_{R \in \mathcal{A}} |\Delta_R f| \), where \( [a, b] = \{ x \in \mathbb{R}^k : a \leq x \leq b \} \), \( a, b \in \mathbb{R}^k, \ k \in \mathbb{N} \). The supremum is taken over all finite collections of \( k \)-dimensional rectangles \( \mathcal{A} = \{ R_i : 1 \leq i \leq m \} \) such that \( \bigcup_{i=1}^{m} R_i = [a, b] \), and the interiors of any two rectangles in \( \mathcal{A} \) are disjoint. Here, if \( R = [c, d] \), a \( k \)-dimensional rectangle contained in \( [a, b] \), then, \( \Delta_R f = \sum_{I \subseteq \{1, 2, ..., k\}} (-1)^{|I|} f(x_I) \), where, \( x_I \) is the vector in \( \mathbb{R}^k \) whose \( i^{th} \) element is given by \( c_i \) if \( i \in I \), or by \( d_i \) if \( i \notin I \), \( f_\emptyset = f(b) \). For instance, if \( k = 2 \) and \( R = [c_1, d_1] \times [c_2, d_2] \) then, \( \Delta_R f = f(d_1, d_2) - f(c_1, d_2) - f(d_1, c_2) + f(c_1, c_2) \).

**Definition 2.4.** The Hardy-Krause variation of a function \( f : [a, b] \to \mathbb{R} \), is given by, \( ||f||_{HK} = \sum_{\emptyset \neq I \subseteq \{1, 2, ..., k\}} |f_I| \), where \( [a, b] = \{ x \in \mathbb{R}^k : a \leq x \leq b \} \), \( a, b \in \mathbb{R}^k, \ k \in \mathbb{N} \). Here, given a non-empty set \( I \subseteq \{1, 2, ..., k\} \), \( f_I \) denotes the real valued function on \( \prod_{i \in I} [a_i, b_i] \) obtained by setting the \( i^{th} \) argument of \( f \) equal to \( b_i \) whenever \( i \notin I \).

When \( k = 1 \), the Hardy-Krause variation is equivalent to Vitali variation and hence the standard definition of total variation.

The following results are discussed for non-degenerate U-statistics of degree 2.
Assume that $F$ is the distribution function of $X_1$. Let the U-statistic with a kernel $\rho$ of degree 2 based on observations $\{X_j, 1 \leq j \leq n\}$ be denoted by $U_n(\rho)$.

$$U_n(\rho) = \frac{2}{n(n-1)} \sum_{1 \leq i < j \leq n} \rho(X_i, X_j) = \theta + \frac{2}{n} \sum_{i=1}^{n} h^{(1)}(X_i) + \frac{2}{n(n-1)} \sum_{1 \leq i < j \leq n} h^{(2)}(X_i, X_j). \quad (2.3)$$

The equality in (2.3) is due to Hoeffding’s decomposition (Hoeffding (1948)), and $h^{(1)}$ and $h^{(2)}$ are given as follows. Define $\theta = \int_{R^2} \rho(x,y) dF(x) dF(y)$.

$$\rho_1(x_1) = \int_{R} \rho(x_1,x_2) dF(x_2), \quad h^{(1)}(x_1) = \rho_1(x_1) - \theta,$$

and $h^{(2)}(x_1,x_2) = \rho(x_1,x_2) - \rho_1(x_1) - \rho_1(x_2) + \theta$.

**Lemma 2.5.** (Garg and Dewan (2015)) Let $P(|X_n| \leq C_1) = 1$ for some $0 < C_1 < \infty$. Let $h^{(2)}(x,y)$ be a degenerate kernel of degree 2 (i.e. $\int_{R} h^{(2)}(x,y) dF(y) = 0$ for all $x \in R$), and $|h^{(2)}(x,y)| \leq M(C_1)$, for some $0 < M(C_1) < \infty$, for all $x,y \in [-C_1,C_1]$. Assume that the density function of $X_1$ is bounded and $h^{(2)}$ is of bounded Hardy-Krause variation and left-continuous. Further, let $\sum_{j=1}^{\infty} Cov(X_1,X_j)^{\gamma} < \infty$, for some $0 < \gamma < 1/6$. Then, as $n \to \infty$,

$$\sum_{1 \leq i < j \leq n} \sum_{1 \leq k \leq n} |E(h^{(2)}(X_i,X_j)h^{(2)}(X_k,X_l))| = O(n^2). \quad (2.4)$$

**Lemma 2.6.** Assume that the conditions of Lemma 2.5 are true. Let $\sigma_1^2 = Var(h^{(1)}(X_1)) < \infty$, and $\sum_{j=1}^{\infty} |\sigma_j| < \infty$, where $\sigma_j = Cov(h^{(1)}(X_1),h^{(1)}(X_{1+j}))$, $j \geq 1$. Then, as $n \to \infty$,

$$Var(U_n) = \frac{4\sigma_U^2}{n} + o\left(\frac{1}{n}\right), \quad \text{where } \sigma_U^2 = \sigma_1^2 + 2 \sum_{j=1}^{\infty} \sigma_j. \quad (2.5)$$

**Lemma 2.7.** Assume that the conditions of Lemma 2.6 are satisfied. Let $\sigma_U^2 > 0$. If there exists a function $\tilde{h}^{(1)}(\cdot)$ such that $h^{(1)} \ll \tilde{h}^{(1)}$ and,

$$\sum_{j=1}^{\infty} Cov(\tilde{h}^{(1)}(X_1),\tilde{h}^{(1)}(X_j)) < \infty.$$

Then,

$$\frac{\sqrt{n}(U_n - \theta)}{2\sigma_U} \overset{d}{\to} N(0,1) \quad \text{as } n \to \infty, \quad (2.6)$$

where $\sigma_U^2$ is defined in (2.5).

**Remark 2.1.** If $h_1 \ll \tilde{h}_1$, then $\sigma_1^2 \leq Var(\tilde{h}_1(X_1))$ and $|\sigma_j| \leq C Cov(\tilde{h}_1(X_1),\tilde{h}_1(X_j))$. If $h^{(1)}$ is monotonic, then $\tilde{h}_1 \equiv h^{(1)}$ and $\{h^{(1)}(X_n), n \geq 1\}$ is a sequence of stationary associated random variables.
Remark 2.2. The condition that the random variables are uniformly bounded is only required to prove the covariance inequality in Lemma 2.5. The covariance inequality can be extended to random variables which are not uniformly bounded by the usual truncation techniques. We illustrate this for the U-statistic estimators for the second and the third L-moments in Theorem 5.1, Section 5.1.

Remark 2.3. The results of Lemmas 2.5−2.7 can be easily extended to non-degenerate U-statistics based on kernels of finite degrees greater than 2.

Remark 2.4. The limiting distribution of U-statistics based on \( \{X_n, n \geq 1\} \) with kernels that are of bounded variation can also be obtained using the results of Beutner and Zähle (2012, 2014), under a different set of assumptions. The approach of Beutner and Zähle (2012) is based on a modified delta method and quasi-Hadamard differentiability, while Beutner and Zähle (2014) proposed a continuous mapping approach. Their results require weak convergence for weighted empirical processes of the underlying random variables. Their techniques cover a large set of dependence structures, like \( \rho \)-mixing, \( \alpha \)-mixing, association, etc. and even long-range dependent sequences.

3 Bootstrap for functions of stationary associated random variables

In the following, we prove consistency of the estimators of distribution function and variance of the sample mean of \( \{g(X_j), 1 \leq j \leq n\} \) with \( g : \mathbb{R} \rightarrow \mathbb{R} \) obtained using Circular Block Bootstrap (CBB). We assume that, there exists \( \tilde{g} : \mathbb{R} \rightarrow \mathbb{R} \), such that \( g \ll \tilde{g} \). Define, \( Y_n = g(X_n) \), and \( \tilde{Y}_n = \tilde{g}(X_n) \), \( n \geq 1 \).

Let \( \Omega_n = \{Y_i, 1 \leq i \leq n\} \) have a common one-dimensional marginal distribution function and \( E(Y_1) = \mu \). Let the following be of interest:

\[
T_n = \sqrt{n}(\bar{Y}_n - \mu), \quad \text{and} \quad G_n(x) = P(T_n \leq x), x \in \mathbb{R}.
\]  

The Circular Block Bootstrap (CBB) method was proposed by Politis and Romano (1992). This method re-samples overlapping and periodically extended blocks of a given length \( \ell \equiv \ell_n \), \( \ell \) is a positive integer) satisfying \( \ell = o(n) \) as \( n \rightarrow \infty \) from \( \{B(1, \ell), \cdots, B(n, \ell)\} \). \( B(i, \ell), i = 1, \cdots, n \), are defined as follows.

\[
B(i, \ell) = (Y_{n,i}, \cdots, Y_{n,i+\ell-1}), \quad \text{where}
\]

\[
Y_{n,i} = Y_i, \quad \text{if} \quad i = 1, \cdots, n,
\]

\[
= Y_j \quad \text{if} \quad j = i - n, \quad i = n + 1, \cdots, n + (\ell - 1)
\]

To obtain the CBB sample, randomly select \( k \) blocks from \( \{B(1, \ell), \cdots, B(n, \ell)\} \) with replacement. The sample size is \( m = k\ell \). Let \( \Omega_m^* = \{Y_{i}^*, 1 \leq i \leq m\} \) denote the CBB sample of size \( m \) from \( \Omega_n \).
Let \( \{B^*(1, \ell), \cdots, B^*(k, \ell)\} \) denote the selected sample of blocks and the elements in \( B^*(j, \ell) \) be denoted as \((Y_{(j-1)\ell+1}^*, \cdots, Y_{j\ell}^*)\), \( j = 1, 2, \cdots, k \).

\[
P_\ast((Y_1^*, \cdots, Y_{n, i}^*)') = \frac{1}{n}, \quad i = 1, \cdots, n,
\]

where \( P_\ast \) denotes the conditional probability given \( \Omega_n \). Note that in CBB equal weights are given to each of the observations \( Y_1, \cdots, Y_n \).

For our calculations, we used \( k = \left[\frac{n}{\ell}\right] \) and \( m = k\ell \).

Let \( E_\ast \) and \( Var_\ast \) respectively denote the conditional expectation and conditional variance, given \( \Omega_n \). Then, the bootstrap version of \( T_n \) is given by,

\[
T_n^\ast = \sqrt{m}(\bar{Y}_m^* - E_\ast\bar{Y}_m^*) = \sqrt{m}(\bar{Y}_m^* - \bar{Y}_n),
\]

as \( E_\ast\bar{Y}_m^* = \bar{Y}_n \) (from Lahiri (2003) (Section 2.7.1, (2.18))). Similarly, the bootstrap estimator for \( G_n(x) \) is \( G_n^\ast(x) = P_\ast(T_n^\ast \leq x) \). Note that, \( \lim_{n \to \infty} Var(T_n) = Var(Y_1) + 2 \sum_{j=2}^{\infty} Cov(Y_1, Y_j) \).

In the following section, we prove that \( Var_\ast(T_n^\ast) \) is a consistent estimator of \( \lim_{n \to \infty} Var(T_n) \), and \( G_n^\ast(x) \) is a consistent estimator of the sampling distribution \( G_n(x), x \in \mathbb{R} \).

### 3.1 Consistency of \( Var_\ast(T_n^\ast) \)

Let,

\[
U_i = \frac{Y_{n,i} + \cdots + Y_{n,i+\ell-1}}{\ell},
\]

be the average of \( B(i, \ell) \), \( i = 1, \cdots, n \). As re-sampled blocks are independent,

\[
Var_\ast(T_n^\ast) = \ell \left[ n^{-1} \sum_{i=1}^{n} U_i^2 - \bar{Y}_n^2 \right]
\]

The proofs of the following are similar to the results of Lahiri (2003) (Section 3.2.1), albeit some minor modifications.

**Theorem 3.1.** Assume that the conditions of Lemma 2.2 are true, and that,

\[
\sum_{j=1}^{\infty} Cov(\bar{Y}_1, \bar{Y}_j)^{1/3} < \infty.
\]

Further, suppose that \( (\sum_{j=1}^{\ell} Y_j - \ell E(Y_1))/\sqrt{\ell} \) has a bounded continuous density for all \( \ell \in \mathbb{N} \). Then,

\[
Var_\ast(T_n^\ast) \overset{p}{\to} \sigma^2 \text{ as } n \to \infty,
\]

where \( \sigma^2 \) is defined in (2.1).

**Proof.** Let \( C \) be a generic positive constant in the sequel. Assume without loss of generality that \( \mu = 0 \). Note that,

\[
E(\ell \bar{Y}_n^2) = O\left(\frac{\ell}{n}\right) \to 0 \text{ as } n \to \infty.
\]
Therefore, we just need to prove that
\[ \ell n^{-1} \sum_{i=1}^{n} U_i^2 \to \sigma^2, \text{ as } n \to \infty, \]  
(3.9)
i.e. we need, for any \( \epsilon > 0 \),
\[ \lim_{n \to \infty} P \left( |\ell n^{-1} \sum_{i=1}^{n} U_i^2 - \sigma^2| > 6\epsilon \right) = 0. \]  
(3.10)
Let \( U_{1i} = \sqrt{\ell} U_i, i = 1, \ldots, n \), and \( N = n - \ell + 1 \) in the following.
\[ P \left( |n^{-1} \sum_{i=1}^{n} U_{1i}^2 - \sigma^2| > 6\epsilon \right) \leq P \left( |n^{-1} \sum_{i=1}^{N} (U_{1i}^2 - \sigma^2)| > 3\epsilon \right) + P \left( |n^{-1} \sum_{i=N+1}^{n} (U_{1i}^2 - \sigma^2)| > 3\epsilon \right) \]
\[ \leq C \left[ P \left( n^{-1} \sum_{i=1}^{N} |V_{in} - E(V_{in})| > \epsilon \right) + P \left( |n^{-1} \sum_{i=1}^{N} W_{in}| > \epsilon \right) \right. \]
\[ + \left. \frac{N}{n} |E(V_{1n}) - \sigma^2| + P \left( |n^{-1} \sum_{i=N+1}^{n} (U_{1i}^2 - \sigma^2)| > 3\epsilon \right) \right], \]  
(3.11)
where \( V_{in} = U_{1i}^2 I(|U_{1i}| < \left( \frac{n}{\ell} \right)^{1/8}) \) and \( W_{in} = U_{1i}^2 - V_{in} \) for \( i = 1, \ldots, N \).
Define, \( \tilde{U}_{1i} = \frac{Y_{n,i+\ell-1} + \cdots + Y_{n,i+1}}{\sqrt{\ell}}, i = 1, 2, \ldots, n \), where \( \{\tilde{Y}_{n,i}, 1 \leq i \leq n\} \) is the periodically extended series of \( \{\tilde{Y}_i, 1 \leq i \leq n\} \). As \( U_{1i} \ll \tilde{U}_{1i}, U_{1i}, i = 1, \ldots, n \) are square integrable. Therefore we get,
\[ P \left( |n^{-1} \sum_{i=N+1}^{n} (U_{1i}^2 - \sigma^2)| > 3\epsilon \right) = O \left( \frac{\ell}{n} \right), \text{ as } n \to \infty. \]  
(3.12)
Next,
\[ \text{Var}(n^{-1} \sum_{i=1}^{N} V_{in}) \leq C \frac{1}{n^3} \left( \frac{n}{\ell} \right)^{1/2} \left( \sum_{i=1}^{N} \text{Var}(\tilde{U}_{1i}) \right)^{1/3} + 2 \sum_{1 \leq i < j \leq N} \text{Cov}(\tilde{U}_{1i}, \tilde{U}_{1j})^{1/3} \]
\[ \leq \frac{1}{n^3} \left( \frac{n}{\ell} \right)^{1/2} O(n\ell) \to 0 \text{ as } n \to \infty. \]  
(3.13)
The last inequality follows using (3.7). Also,
\[ \lim_{n \to \infty} P \left( |n^{-1} \sum_{i=1}^{N} W_{in}| > \epsilon \right) \leq \lim_{n \to \infty} C \frac{E|\sum_{i=1}^{N} W_{in}|}{n\epsilon} \leq \lim_{n \to \infty} C \frac{NE|W_{1n}|}{n\epsilon} \]
\[ \leq \lim_{n \to \infty} C E|U_{11}^2 I(|U_{11}| > \left( \frac{n}{\ell} \right)^{1/8})|. \]
As \( U_{11} = \sqrt{\ell} \tilde{Y}_\ell \), under the conditions of Lemma 2.2, \( \sqrt{n} \tilde{Y}_n \overset{\mathcal{L}}{\to} N(0, \sigma^2) \). By dominated convergence theorem,
\[ \lim_{n \to \infty} E|U_{11}^2 I(|U_{11}| > \left( \frac{n}{\ell} \right)^{1/8})| = 0. \]  
(3.14)
Therefore,

\[
\lim_{n \to \infty} P\left(|n^{-1} \sum_{i=1}^{N} W_{in}| > \epsilon\right) = 0 \quad \text{and} \quad \lim_{n \to \infty} |E(V_{1n}) - \sigma|^2 = 0.
\]  

(3.15)

The result follows using (3.12), (3.13), and (3.15).

\[\square\]

3.2 Consistency of the estimator of bootstrap distribution function

**Theorem 3.2.** Assume that the conditions of Theorem 3.1 are true, then,

\[
\sup_{x \in \mathbb{R}} |G_n^*(x) - G_n(x)| \xrightarrow{p} 0, \quad \text{as} \quad n \to \infty.
\]

(3.16)

**Proof.** Note that, Lemma 2.2, implies

\[
\sup_{x \in \mathbb{R}} |G_n(x) - \Phi(x\sigma)| \to 0 \quad \text{as} \quad n \to \infty.
\]

Hence, it is enough to prove,

\[
\sup_{x \in \mathbb{R}} |G_n^*(x) - \Phi(x\sigma)| \xrightarrow{p} 0, \quad \text{as} \quad n \to \infty.
\]

(3.17)

Let,

\[
U_j^* = Y_{(l-1)j+1} + \cdots + Y_{lj}, \quad j = 1, \ldots, k.
\]

Observe, that given \(\Omega_n\), \(T_n^* = \sum_{j=1}^{k} \sqrt{\ell}(U_j^* - \bar{Y}_n)\) is a sum of independent (not identically distributed) random variables.

Define, for \(a > 0\), \(\hat{\Delta}_n(a) = \ell k^{-1} \sum_{j=1}^{k} E(U_j^* - \bar{Y}_n)^2 I(\sqrt{\ell}(U_j^* - \bar{Y}_n) > 2a)\).

\[
P\left(\hat{\Delta}_n((n/\ell)^{1/4}) > \epsilon\right) \leq \epsilon^{-1} E\hat{\Delta}_n\left((n/\ell)^{1/4}\right)
\]

\[
\leq \epsilon^{-1} E\left[\frac{N}{n}(U_{11} - \sqrt{\ell}\bar{Y}_n)^2 I((U_{11} - \sqrt{\ell}\bar{Y}_n) > 2(n/\ell)^{1/4})\right] + C\frac{\ell}{n}
\]

\[
\leq C\epsilon^{-1} \frac{N}{n} E(U_{11}^2 I(|U_{11}| > (n/\ell)^{1/4}) + \ell E(Y_n^2) + C\frac{\ell}{n} \to 0, \quad \text{as} \quad n \to \infty.
\]

(3.18)

Rest of the proof follows using Lindeberg’s CLT for independent random variables and Theorem 3.1, similarly as in the proof of Lahiri (2003) (Theorem 3.2).

\[\square\]

4 Bootstrap for U-statistics based on a sequence of stationary associated random variables.

The results developed in Section 3 are used to obtain consistent estimators of variance and the distribution function of U-statistics based on a sequence of stationary associated random variables.
4.1 Consistency of estimators of distribution function and limiting variance of U-statistics based on CBB

Let \( \{X^*_i, 1 \leq i \leq k\ell \} \) denote the sample of size \( k\ell \) obtained using CBB from \( \{X_i, 1 \leq i \leq n\} \).

Let the U-statistic with a kernel \( \rho \) of degree 2 based on \( \{X^*_i, 1 \leq i \leq k\ell\} \) be denoted as \( U^*_n(\rho) \), i.e.,

\[
U^*_n(\rho) = \frac{1}{k\ell} \left( \binom{k\ell}{2} \right) \sum_{1 \leq i < j \leq k\ell} \rho(X^*_i, X^*_j) = \theta + \frac{2}{k\ell} \sum_{i=1}^{k\ell} h^{(1)}(X^*_i) + \frac{1}{k\ell} \sum_{1 \leq i < j \leq k\ell} h^{(2)}(X^*_i, X^*_j)
\]

\[
= \theta + \frac{2}{k\ell} \sum_{i=1}^{k\ell} h^{(1)}(X^*_i) + U^*_n(h^{(2)}).
\] (4.1)

The proofs of the following are similar to proofs in Dehling and Wendler (2010).

**Theorem 4.1.** Let the conditions of Lemma 2.5 be true.

\[
Var^*\left[ \sqrt{k\ell} U^*_n(h^{(2)}) \right] \xrightarrow{p} 0, \text{ as } n \to \infty.
\] (4.2)

*Proof.* Let \( C \) be a generic positive constant in the following. Using Lemma 2.5 and following the proof of Dehling and Wendler (2010) (Lemma 3.7), we get,

\[
E\left[ E^*\left( \left( \frac{2}{\sqrt{k\ell(k\ell - 1)}} \sum_{1 \leq i < j \leq k\ell} h^{(2)}(X^*_i, X^*_j) \right)^2 \right) \right] \leq \frac{C}{k\ell(k\ell - 1)^2} \sum_{i_1,i_2,i_3,i_4=1}^n |E(h^{(2)}(X_{i_1}, X_{i_2})h^{(2)}(X_{i_3}, X_{i_4}))| = o(1), \text{ as } n \to \infty.
\] (4.3)

\[\square\]

**Theorem 4.2.** Assume that the conditions of Lemma 2.7 are true, and that,

\[
\sum_{j=1}^{\infty} \text{Cov}(h^{(1)}(X_1), h^{(1)}(X_j))^{1/3} < \infty.
\]

Further, suppose that \( (\sum_{j=1}^{\ell} h^{(1)}(X_j) - \ell\theta)/\sqrt{\ell} \) has a bounded continuous density for all \( \ell \in \mathbb{N} \). Then,

\[
\begin{align*}
|Var^*\left[ \sqrt{k\ell} U^*_n(\rho) \right] - Var\left[ \sqrt{n} U_n(\rho) \right] | \xrightarrow{p} 0, & \text{ as } n \to \infty, \\
\sup_{x \in \mathbb{R}} P^*\left( \sqrt{\ell} \left( U^*_n(\rho) - E^*\left[ U^*_n(\rho) \right] \right) \leq x \right) - \left( \sqrt{n} \left( U_n(\rho) - \theta \right) \leq x \right) | \xrightarrow{p} 0, & \text{ as } n \to \infty.
\end{align*}
\] (4.4, 4.5)

*Proof.* Using the Hoeffding’s decomposition we have,

\[
U^*_n(\rho) = \theta + \frac{2}{k\ell} \sum_{i=1}^{k\ell} h^{(1)}(X^*_i) + U^*_n(h^{(2)}).
\] (4.6)
Using Lemma 2.5,
\[
\left| \text{Var} \left[ \sqrt{n} U_n(\rho) \right] \right| \rightarrow \left| \text{Var} \left[ \frac{2}{n} \sum_{i=1}^{n} h^{(1)}(X_i) \right] \right| \quad \text{as } n \rightarrow \infty. \tag{4.7}
\]

Using Theorem 4.1, we get,
\[
\left| \text{Var}_* \left[ \sqrt{k\ell} U^*_n(\rho) \right] \right| \overset{p}{\rightarrow} \left| \text{Var}_* \left[ \frac{2}{k\ell} \sum_{i=1}^{k\ell} h^{(1)}(X^*_i) \right] \right| \quad \text{as } n \rightarrow \infty. \tag{4.8}
\]

Using Theorem 3.1, we get,
\[
\left| \text{Var}_* \left[ \frac{2}{k\ell} \sum_{i=1}^{k\ell} h^{(1)}(X^*_i) \right] - \text{Var} \left[ \frac{2}{n} \sum_{i=1}^{n} h^{(1)}(X_i) \right] \right| \overset{p}{\rightarrow} 0, \quad \text{as } n \rightarrow \infty. \tag{4.9}
\]

Hence, (4.4) is proved. Finally, as \( n \rightarrow \infty \),
\[
\sup_{x \in \mathbb{R}} P \left( \left. \sqrt{k\ell} \left( U^*_n(\rho) - E_* \left[ U^*_n(\rho) \right] \right) \leq x \right| - P \left( \frac{2}{\sqrt{k\ell}} \sum_{i=1}^{k\ell} \left( h^{(1)}(X^*_i) - E_* \left[ h^{(1)}(X^*_i) \right] \right) \leq x \right) \right) \overset{p}{\rightarrow} 0. \tag{4.10}
\]

Using (4.10), (4.11) and Theorem 3.2, we get (4.5).

\[ \square \]

**Remark 4.1.** These results can be easily extended to non-degenerate U-statistics based on kernels with finite degrees greater than 2.

**Remark 4.2.** The results of Sections 3 – 4 will also hold for moving block and non-overlapping block bootstrap. The proofs will follow similarly.

## 5 Point and Interval estimation for L-moments

L-moments uniquely characterize a distribution function (provided the mean exists). They are more robust to the effect of outliers in data and allow for more reliable inferences about the underlying probability distribution than conventional moments (see, Hosking (1990)). Applications of L-moments can be found in economics, engineering, meteorology, reliability, and hydrology. For examples, see Hosking (1990), Jones and Balakrishnan (2002), Fitzgerald (2005), Yang et. al. (2010), Wang et. al. (2010), among others.

The \( r \)th L-moment, \( \lambda_r \), of a random variable \( W_1 \) with the distribution function \( F \), is a function of the expected order statistics of a random sample of size \( r \) from \( F \). Let \( \{W_i, 1 \leq i \leq r\} \) be a random sample from \( F \). Define, \( W_{1:r} \leq W_{2:r} \leq \cdots \leq W_{r:r} \) as the corresponding ordered sample. Then,
\[
\lambda_r = r^{-1} \sum_{k=0}^{r-1} (-1)^{k} \binom{r-1}{k} E(W_{r-(r-k)}), \quad r = 1, 2, \cdots
\]
In particular, the first three L-moments for $W_1$ are,

\[
\begin{align*}
\lambda_1 &= E(W_1), \\
\lambda_2 &= \frac{1}{2} \left( E(W_{2:2}) - E(W_{1:2}) \right), \\
\lambda_3 &= \frac{1}{3} \left( E(W_{3:3}) - 2E(W_{2:3}) + E(W_{1:3}) \right).
\end{align*}
\]

(5.1)

$\lambda_1$, $\lambda_2$, and $\lambda_3$ can be used for measuring descriptive features (location, scale and skewness, respectively) of $F$.

In this section, we look at point and interval estimation for L-moments, when the underlying sample consists of stationary associated random variables.

5.1 Consistent point estimators for L-moments

Let $\{X_i, 1 \leq i \leq n\}$ be a sample of stationary associated random variables from $F$. Define, $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$ as the corresponding ordered sample. $\lambda_r$ can be estimated by a U-statistic,

\[
\hat{\lambda}_{r,n} = r^{-1} \left( \begin{array}{c} n \cr r \end{array} \right)^{-1} \sum_{1 \leq i_1 \leq \cdots \leq i_r \leq n} \sum_{k=0}^{r-1} \left( \begin{array}{c} r-1 \cr k \end{array} \right) (-1)^k X_{i_r-k:n}, \ r = 1, \cdots, n.
\]

(5.2)

And, in particular,

\[
\begin{align*}
\hat{\lambda}_{1,n} &= n^{-1} \sum_{1 \leq i \leq n} X_{i:n}, \\
\hat{\lambda}_{2,n} &= \left( \begin{array}{c} n \cr 2 \end{array} \right)^{-1} \sum_{1 \leq i_1 < i_2 \leq n} (X_{i_2:n} - X_{i_1:n}), \\
\hat{\lambda}_{3,n} &= \left( \begin{array}{c} n \cr 3 \end{array} \right)^{-1} \sum_{1 \leq i_1 < i_2 < i_3 \leq n} (X_{i_3:n} - 2X_{i_2:n} + X_{i_1:n}).
\end{align*}
\]

(5.3)

$\hat{\lambda}_{1,n}$, is the mean of the sample. Under the conditions of Lemma 2.2, $\hat{\lambda}_{1,n}$ is a weakly consistent estimator of $\lambda_1$. The consistent estimators for the variance and the distribution function of $\hat{\lambda}_{1,n}$, can be obtained using Theorems 3.1 – 3.2 $(g(x) \equiv \tilde{g}(x) = x)$.

$\hat{\lambda}_{r,n}$, $r = 2, 3$ discussed above are U-statistics based on kernels that are functions of bounded Hardy-Krause variation. Under the conditions of Lemma 2.6, $\hat{\lambda}_{r,n}$ are weakly consistent estimators of $\lambda_r$, $r = 2, 3$. The consistent estimators of variance and distribution function of $\hat{\lambda}_{r,n}$, $r = 2, 3$ based on bootstrap can be obtained under the conditions of Theorem 4.2.

The next theorem extends the results of Lemmas 2.5 – 2.7 to non-uniformly bounded random variables for $\hat{\lambda}_{2,n}$ and $\hat{\lambda}_{3,n}$.

\[
\hat{\lambda}_{2,n} = \left( \begin{array}{c} n \cr 2 \end{array} \right)^{-1} \sum_{1 \leq i < j \leq n} \rho(X_i, X_j), \text{ where } \rho(x, y) = |x - y|/2.
\]

11
\[ \hat{\lambda}_{3,n} = \binom{n}{3}^{-1} \sum_{1 \leq i < j < k \leq n} t(X_i, X_j, X_k), \]

where \( t(x, y, w) = (3\max(x, y, w) - 2(x + y + w) + 3\min(x, y, w))/3 \).

**Theorem 5.1.** Let \( \{X_n, n \geq 1\} \) be a sequence of stationary associated random variables, such that, \( E|X_1|^{2+\delta} < \infty \), for some \( \delta > 0 \). Assume that the density function of \( X_1 \) is bounded. Let \( \sum_{j=1}^{\infty} \text{Cov}(X_1, X_j)^{1/2+\delta} < \infty \), for some \( 0 < \gamma < 1/6 \).

1. Let \( h^{(1)} \) and \( h^{(2)} \) be the kernels obtained using the Hoeffding’s decomposition of \( \hat{\lambda}_{2,n} \). Then, as \( n \to \infty \),

\[
\sum_{1 \leq i < j \leq n} \sum_{1 \leq k < j \leq n} |E(h^{(2)}(X_i, X_j)h^{(2)}(X_k, X_l))| = O(n^2).
\]

Further, if \( \sigma_1^2 > 0 \), where \( \sigma_1^2 = \text{Var}(h^{(1)}(X_1)) + 2 \sum_{j=2}^{\infty} \text{Cov}(h^{(1)}(X_1), h^{(1)}(X_j)) \), then, as \( n \to \infty \),

\[
\sqrt{n}(\hat{\lambda}_{2,n} - \lambda_2) \frac{1}{2\sigma_1} \xrightarrow{d} N(0,1).
\]

Also, \( \text{Cov}(h^{(1)}(X_1), h^{(1)}(X_j)) \leq \text{Cov}(X_1, X_j) \), \( j \geq 1 \).

2. Let \( t^{(1)} \), \( t^{(2)} \) and \( t^{(3)} \) be the kernels obtained using Hoeffding’s decomposition for \( \hat{\lambda}_{3,n} \). Then, as \( n \to \infty \),

\[
\sum_{1 \leq i < j \leq n} \sum_{1 \leq k < j \leq n} |E(t^{(2)}(X_i, X_j)t^{(2)}(X_k, X_l))| = O(n^2).
\]

\[
\sum_{1 \leq i < j < k \leq n} \sum_{1 \leq i' < j' < k' \leq n} |E(t^{(3)}(X_i, X_j, X_k)t^{(3)}(X_{i'}, X_{j'}, X_{k'}))| = O(n^5).
\]

Further if \( \sigma_2^2 > 0 \), where \( \sigma_2^2 = \text{Var}(t^{(1)}(X_1)) + 2 \sum_{j=2}^{\infty} \text{Cov}(t^{(1)}(X_1), t^{(1)}(X_j)) \), then, as \( n \to \infty \),

\[
\sqrt{n}(\hat{\lambda}_{3,n} - \lambda_3) \frac{1}{3\sigma_2} \xrightarrow{d} N(0,1).
\]

Also, \( \text{Cov}(t^{(1)}(X_1), t^{(1)}(X_j)) \leq \text{Cov}(X_1, X_j) \gamma^{1/2}, j \geq 1 \).

**Remark 5.1.** Replacing the conditions of Lemmas 2.5 – 2.7 with the conditions of Theorem 5.1, the consistency of \( \hat{\lambda}_{2,n} \) and \( \hat{\lambda}_{3,n} \) and, consistency of the variance and distribution function estimators of \( \hat{\lambda}_{2,n} \) and \( \hat{\lambda}_{3,n} \) based on CBB observations follow for non-uniformly bounded random variables.

The above result can be similarly extended to \( \hat{\lambda}_{r,n}, r = 4, \ldots, n \).

### 5.2 Interval estimation for the L-moments

Interval estimates for the L-moments, when the underlying sample consists of stationary associated random variables, can be obtained using the asymptotic normality or using Circular Block Bootstrap.
\( \hat{\lambda}_{1,n} \) is the sample mean. Its asymptotic normality follows using Lemma 2.2, and the consistency of the estimator of the distribution function using CBB follows from Theorem 3.2 (\( g(x) \equiv \tilde{g}(x) = x \)).

For \( \hat{\lambda}_{2,n} \) and \( \hat{\lambda}_{3,n} \) asymptotic normality follows using Theorem 5.1, while the consistency of the estimator of the distribution function based on CBB follows from Theorem 4.2.

Results for higher order L-moments can be obtained similarly.

5.3 Example - A Marshall-Olkin Log-Logistic process

Many authors (Lewis and McKenzie (1991), Balakrishna and Jacob (2003), Alice and Jose (2005), Jose et al. (2010), among others) have developed and discussed statistical properties of various autoregressive models with minification structures. In the following, we discuss a Marshall-Olkin Log-Logistic process with a minification structure generating a sequence of stationary associated random variables. A case study is also presented.

The survival function of MO Log-Logistic Distribution (\( MO - LLG(\beta, \gamma, p) \)) is,

\[
\bar{F}(x; p, \beta, \gamma, g) = \frac{1}{p(x - g)} + 1, \quad \beta, \gamma, p > 0, x > g.
\]  

(5.4)

When \( p = 1 \), \( \bar{F} \) is survival function of a Log-logistic (\( LLG(\beta, \gamma, g) \)) random variable with shape parameter \( \beta \), location parameter \( g \), and scale parameter \( \gamma \). The statistical properties of \( MO - LLG(\alpha, \beta, \gamma, 0) \) and its application in minification processes have been discussed in Gui (2013).

Following is a Marshall-Olkin Log-Logistic process generating a sequence of stationary associated random \( \{X_n, n \geq 1\} \) with common one-dimensional marginal distribution \( MO - LLG(\alpha, \beta, p, g) \). Assume \( X_0 \) follows \( MO - LLG(\alpha, \beta, p, g) \).

\[
X_n = \begin{cases} 
\epsilon_n \text{ with probability } p, \\
\min(X_{n-1}, \epsilon_n) \text{ with probability } (1-p), n \geq 1,
\end{cases}
\]  

(5.5)

where \( 0 < p < 1 \) and \( \{\epsilon_n, n \geq 1\} \) is a sequence of i.i.d \( LLG(\alpha, \beta, g) \) random variables independent of \( \{X_n, n \geq 0\} \). Observe that,

\[
Cov(X_0, X_n) \leq (1-p)^n Var(X_0), \quad n \geq 1.
\]  

(5.6)

Note that, \( \{X_n, n \geq 1\} \) is a sequence of stationary associated random variables as non-decreasing functions of associated random variables are associated (Esary et al. (1967)).

5.3.1 Point estimation of L-moments and parameters \( \alpha, \beta, \) and \( g \)

The parameters of MO-LLG process can be written as functions of L-moments.

\[
c = \frac{1}{\beta} = \frac{\lambda_3}{\lambda_2}, \quad a = \alpha p^{c} = \lambda_2 \frac{\sin(\pi c)}{\pi c^2}, \quad g = \lambda_1 - \frac{\lambda_2}{\lambda_3}.
\]  

(5.7)
The conditions of Theorem 5.1 are satisfied under $\beta \geq 3$. As $n \to \infty$, $\hat{\lambda}_{r,n} \to \lambda_r$ in $L_2$, $r = 1, 2, 3$. The condition $\sum_{j=1}^{\infty} \text{Cov}(X_1, X_j)^{\gamma \frac{r}{r+1}} < \infty$, for some $0 < \gamma < 1/6$, is satisfied because of (5.6). Hence, when $p$ is known, replacing the L-moments with sample L-moments in the set of equations given in (5.7) provide weakly consistent estimators for $\alpha, \beta$, and $g$.

For all the following simulations we assumed that the value of $p$ is known. We took $\alpha = 0.2$, $\beta = 5$, $g = 2$, and varied the value of $p$. We used the statistical software R (http://www.r-project.org; R Development Core Team (2014)) for our simulations.

Table 5.1 gives the estimates of the first three L-moments and the parameters $\alpha$, $\beta$, and $g$, based on samples of size $n$ ($n = 100, 200, 500$) generated using the model (5.5). The estimates of the first three L-moments are obtained using (5.3). Let $\hat{\alpha}_n$, $\hat{\beta}_n$, and $\hat{g}_n$, denote the estimators of $\alpha$, $\beta$, and $g$ respectively. The following results are based on $N = 5000$ iterations.

<table>
<thead>
<tr>
<th>$p = 0.3$, $\lambda_1 = 2.1680$, $\lambda_2 = 0.0336$, $\lambda_3 = 0.0067$</th>
<th>$n=100$</th>
<th>$n=200$</th>
<th>$n=500$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\lambda}<em>{1,n}$ (Est.M.S.E($\hat{\lambda}</em>{1,n}$))</td>
<td>2.1682(0.00010)</td>
<td>2.1680(0.00005)</td>
<td>2.1681(0.00002)</td>
</tr>
<tr>
<td>$\hat{\lambda}<em>{2,n}$ (Est.M.S.E($\hat{\lambda}</em>{2,n}$))</td>
<td>0.0303(0.00002)</td>
<td>0.0333(0.00001)</td>
<td>0.0335(0.00000)</td>
</tr>
<tr>
<td>$\hat{\lambda}<em>{3,n}$ (Est.M.S.E($\hat{\lambda}</em>{3,n}$))</td>
<td>0.0071(0.00001)</td>
<td>0.0069(0.00001)</td>
<td>0.0068(0.00000)</td>
</tr>
<tr>
<td>$\hat{\alpha}_n$ (Est.M.S.E($\hat{\alpha}_n$))</td>
<td>0.2553(9.32903)</td>
<td>0.2086(0.01180)</td>
<td>0.2023(0.00141)</td>
</tr>
<tr>
<td>$\hat{\beta}_n$ (Est.M.S.E($\hat{\beta}_n$))</td>
<td>7.0618(10667.2)</td>
<td>5.3655(11.7929)</td>
<td>5.1140(1.21931)</td>
</tr>
<tr>
<td>$\hat{g}_n$ (Est.M.S.E($\hat{g}_n$))</td>
<td>1.9431(9.32904)</td>
<td>1.9905(0.01156)</td>
<td>1.9973(0.00132)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$p = 0.5$, $\lambda_1 = 2.1861$, $\lambda_2 = 0.0372$, $\lambda_3 = 0.0074$</th>
<th>$n=100$</th>
<th>$n=200$</th>
<th>$n=500$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\lambda}<em>{1,n}$ (Est.M.S.E($\hat{\lambda}</em>{1,n}$))</td>
<td>2.1861(0.00009)</td>
<td>2.1860(0.00004)</td>
<td>2.1861(0.00002)</td>
</tr>
<tr>
<td>$\hat{\lambda}<em>{2,n}$ (Est.M.S.E($\hat{\lambda}</em>{2,n}$))</td>
<td>0.0370(0.00002)</td>
<td>0.0371(0.00001)</td>
<td>0.0372(0.00001)</td>
</tr>
<tr>
<td>$\hat{\lambda}<em>{3,n}$ (Est.M.S.E($\hat{\lambda}</em>{3,n}$))</td>
<td>0.0077(0.00001)</td>
<td>0.0076(0.00001)</td>
<td>0.0075(0.00000)</td>
</tr>
<tr>
<td>$\hat{\alpha}_n$ (Est.M.S.E($\hat{\alpha}_n$))</td>
<td>0.2396(1.40422)</td>
<td>0.2112(0.00437)</td>
<td>0.2039(0.00117)</td>
</tr>
<tr>
<td>$\hat{\beta}_n$ (Est.M.S.E($\hat{\beta}_n$))</td>
<td>6.3085(1234.42)</td>
<td>5.3920(3.63175)</td>
<td>5.1356(0.93060)</td>
</tr>
<tr>
<td>$\hat{g}_n$ (Est.M.S.E($\hat{g}_n$))</td>
<td>1.9596(1.40345)</td>
<td>1.9883(0.00421)</td>
<td>1.9959(0.00111)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$p = 0.8$, $\lambda_1 = 2.2045$, $\lambda_2 = 0.04089$, $\lambda_3 = 0.0082$</th>
<th>$n=100$</th>
<th>$n=200$</th>
<th>$n=500$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\lambda}<em>{1,n}$ (Est.M.S.E($\hat{\lambda}</em>{1,n}$))</td>
<td>2.2046(0.00008)</td>
<td>2.2044(0.00004)</td>
<td>2.2044(0.00002)</td>
</tr>
<tr>
<td>$\hat{\lambda}<em>{2,n}$ (Est.M.S.E($\hat{\lambda}</em>{2,n}$))</td>
<td>0.0409(0.00002)</td>
<td>0.04085(0.00001)</td>
<td>0.0409(0.00000)</td>
</tr>
<tr>
<td>$\hat{\lambda}<em>{3,n}$ (Est.M.S.E($\hat{\lambda}</em>{3,n}$))</td>
<td>0.0082(0.00001)</td>
<td>0.00409(0.00000)</td>
<td>0.0082(0.00000)</td>
</tr>
<tr>
<td>$\hat{\alpha}_n$ (Est.M.S.E($\hat{\alpha}_n$))</td>
<td>0.2335(0.09336)</td>
<td>0.2134(0.00368)</td>
<td>0.2054(0.00097)</td>
</tr>
<tr>
<td>$\hat{\beta}_n$ (Est.M.S.E($\hat{\beta}_n$))</td>
<td>5.9978(77.7062)</td>
<td>5.4052(2.7661)</td>
<td>5.1613(0.72606)</td>
</tr>
<tr>
<td>$\hat{g}_n$ (Est.M.S.E($\hat{g}_n$))</td>
<td>1.9665(0.09352)</td>
<td>1.9865(0.00534)</td>
<td>1.9946(0.00092)</td>
</tr>
</tbody>
</table>

In the table above, (1) $\hat{\lambda}_{r,n} = \frac{\sum_{i=1}^{N} \hat{\lambda}_{r,n,i}}{N}$, where $\hat{\lambda}_{r,n,i}$ is estimate of $\lambda_i$ in the $i^{th}$ iteration, $i = 1, 2, \ldots, N$, $r = 1, 2, 3$.

(2) Est. M.S.E($\hat{\lambda}_{r,n}$) = $\frac{\sum_{i=1}^{N} (\hat{\lambda}_{r,n,i} - \hat{\lambda}_{r,n})^2}{N}$, $r = 1, 2, 3$.

Similarly, the values for $\hat{\alpha}_n$, $\hat{\beta}_n$, $\hat{g}_n$, Est.M.S.E($\hat{\alpha}_n$), Est.M.S.E($\hat{\beta}_n$), and Est.M.S.E($\hat{g}_n$) were obtained.

**Observations**

(i) Estimation of $\lambda_1$, $\lambda_2$ and $\lambda_3$: For a fixed set of values of the parameters, as the sample size increases, $\hat{\lambda}_{r,n}$, $r = 1, 2, 3$ become closer to the true values and the corresponding estimated
M.S.Es reduce, i.e. the sample L-moments converge to the true values. As \( p \) becomes closer to 1, this convergence becomes faster.

(ii) \textit{Estimation of }\( \alpha, \beta \text{ and } g \): As \( n \) increases, \( \bar{\hat{\alpha}}_n, \bar{\hat{\beta}}_n, \bar{\hat{g}}_n \) become closer to the true values and the corresponding estimated M.S.Es reduce. As \( p \) becomes closer to 1, this convergence becomes faster.

(iii) The convergences for the parameters are slower than the convergences for the L-moments. In particular, for the set of parameters chosen, it can be seen that larger sample sizes (\( n \geq 200 \)) are needed for a viable consistent estimator of \( \beta \).

5.3.2 Interval estimation for the L-moments

Table 5.2 gives a comparison of the lower tail empirical probabilities obtained using the estimates of the percentiles of \( \hat{\lambda}_{2,n} \) using normal distribution approximation (Theorem 5.1) and the bootstrap technique (\( (4.4) \) and \( (4.5) \)). The results are based on a \( N = 5000 \) iterations of samples of size \( n \). For each iteration, the samples were drawn 1000 times for the bootstrap results. We took the block length \( \ell_n = \lceil n^{1/3} \rceil \), \( n = 20, 50, 100, 200 \). The results obtained using normal distribution approximations are given in the parenthesis (\( . \)).

<table>
<thead>
<tr>
<th>( p = 0.3 )</th>
<th>( n=20 )</th>
<th>( n=50 )</th>
<th>( n=100 )</th>
<th>( n=200 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>EP for the 95th percentile</td>
<td>0.9316(0.9642)</td>
<td>0.9424(0.9608)</td>
<td>0.9464(0.9596)</td>
<td>0.9536(0.961)</td>
</tr>
<tr>
<td>EP for the 50th percentile</td>
<td>0.5792(0.6568)</td>
<td>0.5376(0.5998)</td>
<td>0.5068(0.5616)</td>
<td>0.4986(0.5522)</td>
</tr>
<tr>
<td>EP for the 5th percentile</td>
<td>0.2196(0.2984)</td>
<td>0.1624(0.2036)</td>
<td>0.1320(0.1606)</td>
<td>0.1076(0.1206)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( p = 0.5 )</th>
<th>( n=20 )</th>
<th>( n=50 )</th>
<th>( n=100 )</th>
<th>( n=200 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>EP for the 95th percentile</td>
<td>0.9272(0.9606)</td>
<td>0.9538(0.9688)</td>
<td>0.9568(0.9658)</td>
<td>0.9610(0.9654)</td>
</tr>
<tr>
<td>EP for the 50th percentile</td>
<td>0.5394(0.6128)</td>
<td>0.5068(0.5748)</td>
<td>0.4930(0.5492)</td>
<td>0.4930(0.534)</td>
</tr>
<tr>
<td>EP for the 5th percentile</td>
<td>0.1942(0.2628)</td>
<td>0.1456(0.1806)</td>
<td>0.1206(0.1406)</td>
<td>0.1020(0.1122)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( p = 0.8 )</th>
<th>( n=20 )</th>
<th>( n=50 )</th>
<th>( n=100 )</th>
<th>( n=200 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>EP for the 95th percentile</td>
<td>0.9224(0.9564)</td>
<td>0.9594(0.9692)</td>
<td>0.9664(0.9734)</td>
<td>0.969(0.968)</td>
</tr>
<tr>
<td>EP for the 50th percentile</td>
<td>0.5010(0.5784)</td>
<td>0.4934(0.5464)</td>
<td>0.4724(0.5244)</td>
<td>0.4920(0.5258)</td>
</tr>
<tr>
<td>EP for the 5th percentile</td>
<td>0.1776(0.2406)</td>
<td>0.1370(0.1636)</td>
<td>0.1134(0.1232)</td>
<td>0.1002(0.1058)</td>
</tr>
</tbody>
</table>

To apply Theorem 5.1, an estimator for \( \sigma_1^2 \) is needed. We obtained a consistent estimator by using Theorem 3.1. In general, \( \hat{h}^{(1)} \) would not be known. We used a consistent point estimator \( \hat{h}_n^{(1)} \) of \( h_n^{(1)}(x) = \sum_{j=1}^{n} |X_j - x|/(2n) \). Taking \( Y_i = \hat{h}_n^{(1)}(X_i) \) in \( (3.6) \), we get a consistent estimator for \( \sigma_1^2 \).

Observations

(i) In general, the empirical probabilities obtained using cut-offs from bootstrap are closer to the expected values than the empirical coverage probabilities obtained using the cut-offs from the normal distribution approximation. The bootstrap seems to provide a better estimate of the percentiles than the normal approximation.
(ii) Larger sample sizes are needed to obtain viable estimates for the 5th percentile using both bootstrap and normal distribution approximation.

(iii) The value of $p$ also affects the cut-offs obtained from both the bootstrap as well as the normal distribution approximation. In general, the values of empirical probabilities seem to get closer the expected values as the value of $p$ increases.

Table 5.3 gives the empirical coverage probabilities for the 95% confidence intervals (CIs) using the bootstrap technique for the first three L-moments. The results are based on a $N = 2000$ iterations of samples of size $n$. For each iteration, the samples were drawn 1000 times for the bootstrap results. We took the block length $\ell_n = \lceil n^{1/3} \rceil$, $n = 50, 100, 200, 400$.

It can be seen that as the sample size increases, the empirical coverage probabilities become closer to 0.95.

<table>
<thead>
<tr>
<th>$p$ = 0.3</th>
<th>$n$=50</th>
<th>$n$=100</th>
<th>$n$=200</th>
<th>$n$=400</th>
</tr>
</thead>
<tbody>
<tr>
<td>empirical CP for $\lambda_1$</td>
<td>0.8535</td>
<td>0.8845</td>
<td>0.8995</td>
<td>0.9220</td>
</tr>
<tr>
<td>empirical CP for $\lambda_2$</td>
<td>0.8345</td>
<td>0.8715</td>
<td>0.8960</td>
<td>0.9190</td>
</tr>
<tr>
<td>empirical CP for $\lambda_3$</td>
<td>0.8445</td>
<td>0.8820</td>
<td>0.9020</td>
<td>0.9170</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$p$ = 0.5</th>
<th>$n$=50</th>
<th>$n$=100</th>
<th>$n$=200</th>
<th>$n$=400</th>
</tr>
</thead>
<tbody>
<tr>
<td>empirical CP for $\lambda_1$</td>
<td>0.893</td>
<td>0.9135</td>
<td>0.9325</td>
<td>0.9385</td>
</tr>
<tr>
<td>empirical CP for $\lambda_2$</td>
<td>0.8675</td>
<td>0.8965</td>
<td>0.9135</td>
<td>0.9250</td>
</tr>
<tr>
<td>empirical CP for $\lambda_3$</td>
<td>0.8200</td>
<td>0.8785</td>
<td>0.904</td>
<td>0.9110</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$p$ = 0.8</th>
<th>$n$=50</th>
<th>$n$=100</th>
<th>$n$=200</th>
<th>$n$=400</th>
</tr>
</thead>
<tbody>
<tr>
<td>empirical CP for $\lambda_1$</td>
<td>0.9030</td>
<td>0.919</td>
<td>0.9265</td>
<td>0.9285</td>
</tr>
<tr>
<td>empirical CP for $\lambda_2$</td>
<td>0.878</td>
<td>0.9025</td>
<td>0.9155</td>
<td>0.9245</td>
</tr>
<tr>
<td>empirical CP for $\lambda_3$</td>
<td>0.8215</td>
<td>0.855</td>
<td>0.897</td>
<td>0.9100</td>
</tr>
</tbody>
</table>

**Remark 5.2.** Though, for the example considered in Section 5.3 we have assumed a specific structure for the minification process, the results of Sections 5.3.1 and 5.3.2 can be easily extended to other stationary autoregressive processes with minification structures.

### 5.3.3 Case Study

We fit the discussed MO-LLG process to the dataset consisting of the annual peak flows (in $m^3/s$) of The Thames at Kingston for the years 1883 – 2014 (Source: The UK National River Flow Archive (NRFA), http://nrfa.ceh.ac.uk/data/station/peakflow/39001). The dataset consists of 132 values.

We fitted the model with $p = 0.9$. Using the results of Section 5.3.1, estimates of first three L-moments and the parameters are: $(\hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3) = (327.6378, 64.0396, 8.6281)$, and $(\hat{\alpha}, \hat{\beta}, \hat{g}) = (467.8424, 7.4222, -147.6764)$.

Using (5.4), an estimate of the median annual peak flow is $\hat{\alpha}p^{1/\beta} + \hat{g} = 313.5717$.

Figure 1 gives the Q-Q plot and the plot of the empirical cdf of the observed and the simulated data. It can be seen that the simulated data provides a good fit to the observed data. The parameter $g$ can be taken as an estimate of the minimum of the annual peak flows, but it
is observed in literature (for example, see Fitzgerald (2005)) that reasonable estimates of the quantiles can be obtained even with a negative values of $g$, as is in the present case.

Using the results of Section 5.3.2, the 95% confidence intervals for the first three L-moments were obtained as: for $\lambda_1$: $(312.3809, 343.6756)$, for $\lambda_2$: $(54.9795, 73.7574)$, and for $\lambda_3$: $(2.1364, 14.8681)$. Using the results of Section 3, the 95% confidence interval for the probability that the annual peak flow would be less than the estimated median was obtained as $(0.3936, 0.5616)$, and the 95% confidence interval for the probability that the annual peak flow would be less than $\hat{\lambda}_1$, i.e. the estimated mean, was obtained as $(0.4835, 0.6286)$. These intervals are based on 1000 bootstrap samples.

6 Discussions and intended future work

In this paper, we have discussed the consistency of Circular Block Bootstrap for functions of stationary associated random variables. We also proved the consistency of the estimators of variance and distribution function of U-statistics based on Circular plug-in bootstrap. As applications, interval estimators for L-moments of a stationary sequence of associated random variables are discussed. We have also shown that the U-statistic estimators of L-moments are weakly consistent. To illustrate the use of the theory discussed, we obtain the point estimators and confidence intervals of L-moments of a stationary autoregressive process with a minification structure. Simulations indicate that the point estimates converge faster if the underlying random variables are “almost independent”. A comparison between the estimates of the percentiles of the U-statistics generated via the bootstrap and normal distribution show that, in general, the former provides a better estimate.

We have not discussed the choice of optimal block lengths for the estimators using CBB. Lahiri et. al. (2007) suggested a nonparametric plug-in principle based on the Jackknife-After-Bootstrap (JAB) method to obtain an optimal choice of the block length. They established the
consistency of this method for strongly mixing processes. Extension of this method to stationary associated random variables are under preparation.

References


