isid/ms/2016/09
September 22, 2016
http://www.isid.ac.in/~statmath/index.php?module=Preprint

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 K*ü*nsch(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

Wendler (2012) discussed bootstrap for U-statistics based on mixing processes. Leucht (2012) showed consistency of general bootstrap methods for degenerate U-statistics based on τ - dependent random variables. Dependent multiplier bootstraps for non degenerate U-statistics under mixing conditions was discussed by B$\ddot{u}$cher and Kojadinovic (2016). Sharipov et. al. (2016) suggested an alternative approach of getting a bootstrap version of U-statistics based on random variables satisfying strong mixing and absolute regularity conditions, by drawing with replacement from U-statistics calculated on subsamples.

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, $h_2 : \mathbb{R}^n \to \mathbb{R}$, we have,*

$$Cov(h_1(X_1, \ldots, X_n), h_2(X_1, \ldots, X_n)) \geq 0$$

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 $X_n = a_0\epsilon_n + ... + a_q\epsilon_{n-q}$, ϵ_n are independent random variables and $a_0, ..., a_q$ 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 $\{g(X_n), n \geq 1\}$ $g : \mathbb{R} \to \mathbb{R}$ 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,*

$$\sigma^2 = Var(Y_1) + 2\sum_{j=2}^{\infty} Cov(Y_1, Y_j). \tag{2.1}$$

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

$$\frac{1}{\sqrt{n}\sigma}\sum_{j=1}^{n}(Y_j - E(Y_j)) \xrightarrow{\mathcal{L}} N(0,1) \quad as \quad n \to \infty. \tag{2.2}$$

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 $||f||_V = sup \sum_{R\in\mathbb{A}} |\Delta_R f|$, where $[a,b] = \{\boldsymbol{x} \in \mathbb{R}^k : a \leq \boldsymbol{x} \leq b\}$, $a, b \in \mathbb{R}^k$, $k \in \mathbb{N}$. The supremum is taken over all finite collections of k-dimensional rectangles $\mathbb{A} = \{R_i : 1 \leq i \leq m\}$ such that $\cup_{i=1}^m R_i = [a,b]$, and the interiors of any two rectangles in $\mathbb{A}$ are disjoint. Here, if $R = [c,d]$, a k-dimensional rectangle contained in $[a,b]$, then, $\Delta_R f = \sum_{I\subseteq\{1,2,\ldots,k\}} (-1)^{|I|} f(x_I)$, 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, $||f||_{HK} = \sum_{\emptyset \neq I \subseteq \{1,\ldots,k\}} ||f_I||_V$, where $[a,b] = \{\boldsymbol{x} \in \mathbb{R}^k : a \leq \boldsymbol{x} \leq b\}$, $a, b \in \mathbb{R}^k$, $k \in \mathbb{N}$. 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 ρ of degree 2 based on observations $\{X_j, 1 \le j \le n\}$ be denoted by $U_n(\rho)$.

$$\begin{aligned} U_n(\rho) &= \frac{2}{n(n-1)} \sum_{1\le i<j\le n} \rho(X_i, X_j) \\ &= \theta + \frac{2}{n}\sum_{i=1}^{n} h^{(1)}(X_i) + \frac{2}{n(n-1)} \sum_{1\le i<j\le n} h^{(2)}(X_i, X_j). \end{aligned} \tag{2.3}$$

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

$$\begin{aligned} &\rho_1(x_1) = \int_{\mathbb{R}} \rho(x_1,x_2)\ dF(x_2),\ \ h^{(1)}(x_1) = \rho_1(x_1) - \theta, \\ &\text{and}\ \ h^{(2)}(x_1,x_2) = \rho(x_1,x_2) - \rho_1(x_1) - \rho_1(x_2) + \theta. \end{aligned}$$

Lemma 2.5. *(Garg and Dewan (2015)) Let $P(|X_n| \le C_1) = 1$ for some $0 < C_1 < \infty$. Let $h^{(2)}(x,y)$ be a degenerate kernel of degree 2 (i.e. $\int_{\mathbb{R}} h^{(2)}(x,y)dF(y) = 0$ for all $x \in \mathbb{R}$), and $|h^{(2)}(x,y)| \le M(C_1)$, 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 $n\to\infty$,*

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

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

$$Var(U_n) = \frac{4\sigma_U^2}{n} + o(\frac{1}{n}),\ \ \textit{where}\ \sigma_U^2 = \sigma_1^2 + 2\sum_{j=1}^{\infty}\sigma_{1j}. \tag{2.5}$$

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,

$$\frac{\sqrt{n}(U_n-\theta)}{2\sigma_U} \xrightarrow{\mathcal{L}} N(0,1)\ \ \textit{as}\ n\to\infty, \tag{2.6}$$

where σ_U^2 is defined in (2.5).

Remark 2.1. *If $h_1 \ll \tilde{h}_1$, then $\sigma_1^2 \le Var(\tilde{h}_1(X_1))$ and $|\sigma_{1j}| \le C\,Cov(\tilde{h}_1(X_1), \tilde{h}_1(X_j))$. If $h^{(1)}$ is monotonic, then $\tilde{h}_1 \equiv h^{(1)}$ and $\{h^{(1)}(X_n), n\ge 1\}$ 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 ρ-mixing, α-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 $\{g(X_j), 1 \leq j \leq n\}$, $g : \mathbb{R} \to \mathbb{R}$ obtained using Circular Block Bootstrap (CBB). We assume that, there exists $\tilde{g} : \mathbb{R} \to \mathbb{R}$, 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:

$$T_n = \sqrt{n}(\bar{Y}_n - \mu), \text{ and } G_n(x) = P(T_n \leq x), x \in \mathbb{R}. \tag{3.1}$$

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 $\ell \equiv \ell_n$, (ℓ is a positive integer) satisfying $\ell = o(n)$ as $n \to \infty$ from $\{B(1,\ell), \cdots, B(n,\ell)\}$. $B(i,\ell)$, $i = 1, \cdots, n$, are defined as follows.

$$B(i,\ell) = (Y_{n,i}, \cdots, Y_{n,i+\ell-1}), \text{ where} \tag{3.2}$$

$$\begin{aligned} Y_{n,i} &= Y_i, \text{ if } i = 1, \cdots, n, \\ &= Y_j \quad \text{if } j = i - n, \ i = n+1, \cdots, n + (\ell - 1) \end{aligned}$$

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 $\Omega_m^\star = \{Y_i^\star, 1 \leq i \leq m\}$ denote the CBB sample of size m from Ω_n.

Let $\{B^\star(1,\ell),\cdots,B^\star(k,\ell)\}$ denote the selected sample of blocks and the elements in $B^\star(j,\ell)$ be denoted as $(Y^\star_{(j-1)\ell+1},\cdots,Y^\star_{j\ell})$, $j=1,2,\cdots,k$.

$$P_\star((Y_1^\star,\cdots,Y_\ell^\star)'=(Y_{n,i},\cdots,Y_{n,i+\ell-1})')=\frac{1}{n},\ \ i=1,\cdots,n, \tag{3.3}$$

where $P_\star$ denotes the conditional probability given Ω_n. Note that in CBB equal weights are given to each of the observations $Y_1,\cdots,Y_n$.

For our calculations, we used $k=[\frac{n}{\ell}]$ and $m=k\ell$.

Let $E_\star$ and $Var_\star$ respectively denote the conditional expectation and conditional variance, given Ω_n. Then, the bootstrap version of T_n is given by,

$$T_n^\star=\sqrt{m}(\bar{Y}_m^\star-E_\star\bar{Y}_m^\star)=\sqrt{m}(\bar{Y}_m^\star-\bar{Y}_n), \tag{3.4}$$

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

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

3.1 Consistency of $Var_\star(T_n^\star)$

Let,

$$U_i=\frac{Y_{n,i}+\cdots+Y_{n,i+\ell-1}}{\ell}, \tag{3.5}$$

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

$$Var_\star(T_n^\star)=\ell\Big[n^{-1}\sum_{i=1}^{n}U_i^2-\bar{Y}_n^2\Big] \tag{3.6}$$

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,*

$$\sum_{j=1}^{\infty}Cov(\tilde{Y}_1,\tilde{Y}_j)^{1/3}<\infty. \tag{3.7}$$

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,*

$$Var_\star(T_n^\star)\xrightarrow{p}\sigma^2 \ \text{as}\ n\to\infty, \tag{3.8}$$

where σ^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,

$$E(\ell\bar{Y}_n^2)=O(\frac{\ell}{n})\to 0 \text{ as } n\to\infty.$$

Therefore, we just need to prove that

$$\ell n^{-1}\sum_{i=1}^{n} U_i^2 \xrightarrow{p} \sigma^2, \text{ as } n\to\infty, \tag{3.9}$$

i.e. we need, for any $\epsilon > 0$,

$$\lim_{n\to\infty} P\Big(|\ell n^{-1}\sum_{i=1}^{n} U_i^2 - \sigma^2|> 6\epsilon\Big) = 0. \tag{3.10}$$

Let $U_{1i} = \sqrt{\ell}U_i$, $i = 1, \cdots, n$, and $N = n - \ell + 1$ in the following.

$$\begin{aligned}
P\Big(|n^{-1}\sum_{i=1}^{n} U_{1i}^2 - \sigma^2|> 6\epsilon\Big) &\le P\Big(|n^{-1}\sum_{i=1}^{N}(U_{1i}^2 - \sigma^2)|> 3\epsilon\Big) + P\Big(|n^{-1}\sum_{i=N+1}^{n}(U_{1i}^2 - \sigma^2)|> 3\epsilon\Big)\\
&\le C\Big[P\Big(|n^{-1}\sum_{i=1}^{N}(V_{in} - E(V_{in})\,)|> \epsilon\Big) + P\Big(|n^{-1}\sum_{i=1}^{N} W_{in}|> \epsilon\Big)\\
&\quad + \frac{N}{n}|E(V_{1n}) - \sigma^2|+ P\Big(|n^{-1}\sum_{i=N+1}^{n}(U_{1i}^2 - \sigma^2)|> 3\epsilon\Big)\Big],
\end{aligned} \tag{3.11}$$

where $V_{in} = U_{1i}^2 I(|U_{1i}| < (\frac{n}{\ell})^{1/8})$ and $W_{in} = U_{1i}^2 - V_{in}$ for $i = 1, \cdots, N$.

Define, $\tilde{U}_{1i} = \frac{\tilde{Y}_{n,i}+\cdots+\tilde{Y}_{n,i+\ell-1}}{\sqrt{\ell}}$, $i = 1, 2, \cdots, n$, where $\{\tilde{Y}_{n,i}, 1 \le i \le n\}$ is the periodically extended series of $\{\tilde{Y}_i, 1 \le i \le n\}$. As $U_{1i} \ll \tilde{U}_{1i}$, U_{1i}, $i = 1, \cdots, n$ are square integrable. Therefore we get,

$$P\Big(|n^{-1}\sum_{i=N+1}^{n}(U_{1i}^2 - \sigma^2)|> 3\epsilon\Big) = O(\frac{\ell}{n}), \text{ as } n\to\infty. \tag{3.12}$$

Next,

$$\begin{aligned}
Var(n^{-1}\sum_{i=1}^{N} V_{in}) &\le C\frac{1}{n^2}(\frac{n}{\ell})^{1/2}\Big(\sum_{i=1}^{N} Var(\tilde{U}_{1i})^{1/3} + 2\sum_{1\le i<j\le N} Cov(\tilde{U}_{1i}, \tilde{U}_{1j})^{1/3}\Big)\\
&\le \frac{1}{n^2}(\frac{n}{\ell})^{1/2}O(n\ell) \to 0 \text{ as } \ n\to\infty.
\end{aligned} \tag{3.13}$$

The last inequality follows using (3.7). Also,

$$\begin{aligned}
\lim_{n\to\infty} P\Big(|n^{-1}\sum_{i=1}^{N} W_{in}|> \epsilon\Big) &\le \lim_{n\to\infty} C\frac{E|\sum_{i=1}^{N} W_{in}|}{n\epsilon} \le \lim_{n\to\infty} C\frac{NE|W_{1n}|}{n\epsilon}\\
&\le \lim_{n\to\infty} CE|U_{11}^2 I(|U_{11}|> (\frac{n}{\ell})^{1/8})|.
\end{aligned}$$

As $U_{11} = \sqrt{\ell}\bar{Y}_\ell$, under the conditions of Lemma 2.2, $\sqrt{n}\bar{Y}_n \xrightarrow{\mathcal{L}} N(0, \sigma^2)$. By dominated convergence theorem,

$$\lim_{n\to\infty} E|U_{11}^2 I(|U_{11}|> (\frac{n}{\ell})^{1/8})|= 0. \tag{3.14}$$

Therefore,

$$\underset{n\to\infty}{lim} P\Big(|n^{-1}\sum_{i=1}^{N} W_{in}|> \epsilon\Big) = 0 \text{ and } \underset{n\to\infty}{lim}|E(V_{1n}) - \sigma^2|= 0. \tag{3.15}$$

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

3.2 Consistency of the estimator of bootstrap distribution function

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

$$\underset{x\in R}{sup}\Big|G_n^\star(x) - G_n(x)\Big| \xrightarrow{p} 0, \text{ as } n\to\infty. \tag{3.16}$$

Proof. Note that, Lemma 2.2, implies

$$\underset{x\in R}{sup}|G_n(x) - \Phi(x\sigma)|\to 0 \text{ as } n\to\infty.$$

Hence, it is enough to prove,

$$\underset{x\in R}{sup}|G_n^\star(x) - \Phi(x\sigma)|\xrightarrow{p} 0, \text{ as } n\to\infty. \tag{3.17}$$

Let,

$$U_j^\star = \frac{Y^\star_{(\ell-1)j+1} + \cdots + Y^\star_{\ell j}}{\ell}, \; j = 1,\cdots,k.$$

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

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

$$\begin{aligned}
&P\Big(\hat{\Delta}_n((n/\ell)^{1/4}) > \epsilon\Big) \le \epsilon^{-1}E\hat{\Delta}_n\Big((n/\ell)^{1/4}\Big)\\
&\le \epsilon^{-1}E\Big[\frac{N}{n}(U_{11} - \sqrt{\ell}\bar{Y}_n)^2 I((U_{11} - \sqrt{\ell}\bar{Y}_n) > 2(n/\ell)^{1/4})\Big] + C\frac{\ell}{n}\\
&\le C\epsilon^{-1}\frac{N}{n}\Big[E(U_{11}^2 I(|U_{11}|> (n/\ell)^{1/4}) + \ell E(\bar{Y}_n^2)\Big] + C\frac{\ell}{n} \to 0, \text{ as } n\to\infty.
\end{aligned} \tag{3.18}$$

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). □

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 $\{X_i^\star, 1 \le i \le k\ell\}$ denote the sample of size $k\ell$ obtained using CBB from $\{X_i, 1 \le i \le n\}$. Let the U-statistic with a kernel ρ of degree 2 based on $\{X_i^\star, 1 \le i \le k\ell\}$ be denoted as $U_n^\star(\rho)$, i.e.,

$$\begin{aligned} U_n^\star(\rho) &= \frac{1}{\binom{k\ell}{2}} \sum_{1\le i<j\le k\ell} \rho(X_i^\star, X_j^\star) = \theta + \frac{2}{k\ell}\sum_{i=1}^{k\ell} h^{(1)}(X_i^\star) + \frac{1}{\binom{k\ell}{2}} \sum_{1\le i<j\le k\ell} h^{(2)}(X_i^\star, X_j^\star) \\ &= \theta + \frac{2}{k\ell}\sum_{i=1}^{k\ell} h^{(1)}(X_i^\star) + U_n^\star(h^{(2)}). \end{aligned} \tag{4.1}$$

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.*

$$Var_\star\Big[\sqrt{k\ell}\,U_n^\star(h^{(2)})\Big] \xrightarrow{p} 0, \ \ as\ n \to \infty. \tag{4.2}$$

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,

$$\begin{aligned} &E\Big[E_\star\Big[\Big(\frac{2}{\sqrt{k\ell}(k\ell-1)}\sum_{1\le i<j\le k\ell} h^{(2)}(X_i^\star, X_j^\star)\Big)^2\Big]\Big] \\ &\le \frac{C}{k\ell(k\ell-1)^2}\sum_{i_1,i_2,i_3,i_4=1}^{n} |E(h^{(2)}(X_{i_1}, X_{i_2})h^{(2)}(X_{i_3}, X_{i_4}))| = o(1), \ \text{as } n \to \infty. \end{aligned} \tag{4.3}$$

□

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

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

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,*

$$\Big|Var_\star\Big[\sqrt{k\ell}\,U_n^\star(\rho)\Big] - Var\Big[\sqrt{n}\,U_n(\rho)\Big]\Big| \xrightarrow{p} 0, \ \ as\ n\to\infty, \tag{4.4}$$

$$\sup_{x\in R}\Big|P_\star\Big(\sqrt{k\ell}\Big(U_n^\star(\rho) - E_\star\Big[U_n^\star(\rho)\Big]\Big) \le x\Big) - P\Big(\sqrt{n}\Big(U_n(\rho)-\theta\Big) \le x\Big)\Big| \xrightarrow{p} 0, \ \ as\ n\to\infty. \tag{4.5}$$

Proof. Using the Hoeffding's decomposition we have,

$$U_n^\star(\rho) = \theta + \frac{2}{k\ell}\sum_{i=1}^{k\ell} h^{(1)}(X_i^\star) + U_n^\star(h^{(2)}). \tag{4.6}$$

Using Lemma 2.5,

$$\Big| Var\Big[\sqrt{n}U_n(\rho)\Big] \to Var\Big[\frac{2}{n}\sum_{i=1}^{n} h^{(1)}(X_i)\Big]\Big| \text{ as } n \to \infty. \tag{4.7}$$

Using Theorem 4.1, we get,

$$\Big| Var_\star\Big[\sqrt{k\ell}U_n^\star(\rho)\Big] \xrightarrow{p} Var_\star\Big[\frac{2}{k\ell}\sum_{i=1}^{k\ell} h^{(1)}(X_i^\star)\Big]\Big| \text{ as } n \to \infty \tag{4.8}$$

Using Theorem 3.1, we get,

$$\Big| Var_\star\Big[\frac{2}{k\ell}\sum_{i=1}^{k\ell} h^{(1)}(X_i^\star)\Big] - Var\Big[\frac{2}{n}\sum_{i=1}^{n} h^{(1)}(X_i)\Big]\Big| \xrightarrow{p} 0, \text{ as } n \to \infty. \tag{4.9}$$

Hence, (4.4) is proved. Finally, as $n \to \infty$,

$$\sup_{x\in R}\Big| P_\star\Big(\sqrt{k\ell}\Big(U_n^\star(\rho) - E_\star\Big[U_n^\star(\rho)\Big]\Big) \le x\Big) - P_\star\Big(\frac{2}{\sqrt{k\ell}}\sum_{i=1}^{k\ell}\Big(h^{(1)}(X_i^\star) - E_\star\Big[h^{(1)}(X_i^\star)\Big]\Big) \le x\Big)\Big| \xrightarrow{p} 0. \tag{4.10}$$

$$\sup_{x\in R}\Big| P\Big(\sqrt{n}\Big(U_n(\rho) - E\Big[U_n(\rho)\Big]\Big) \le x\Big) - P\Big(\frac{2}{\sqrt{n}}\sum_{i=1}^{n} h^{(1)}(X_i) \le x\Big)\Big| \to 0. \tag{4.11}$$

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

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, λ_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 $\{W_i, 1 \le i \le r\}$ be a random sample from F. Define, $W_{1:r} \le W_{2:r} \le \cdots \le W_{r:r}$ as the corresponding ordered sample. Then,

$$\lambda_r = r^{-1}\sum_{k=0}^{r-1}(-1)^k\binom{r-1}{k}E(W_{r-k:r}),\ r = 1, 2, \cdots$$

In particular, the first three L-moments for W_1 are,

$$\lambda_1 = E(W_1),$$
$$\lambda_2 = \frac{1}{2}\Big(E(W_{2:2}) - E(W_{1:2})\Big),$$
$$\lambda_3 = \frac{1}{3}\Big(E(W_{3:3}) - 2E(W_{2:3}) + E(W_{1:3})\Big). \tag{5.1}$$

λ_1, λ_2, and λ_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. λ_r can be estimated by a U-statistic,

$$\hat{\lambda}_{r,n} = r^{-1}\binom{n}{r}^{-1} \sum_{1\leq i_1 \leq \cdots \leq i_r \leq n} \sum_{k=0}^{r-1} \binom{r-1}{k}(-1)^k X_{i_{r-k}:n}, \ r = 1, \cdots, n. \tag{5.2}$$

And, in particular,

$$\hat{\lambda}_{1,n} = n^{-1} \sum_{1\leq i\leq n} X_{i:n},$$
$$\hat{\lambda}_{2,n} = \binom{n}{2}^{-1} \frac{1}{2} \sum_{1\leq i_1 < i_2 \leq n} (X_{i_2:n} - X_{i_1:n}),$$
$$\hat{\lambda}_{3,n} = \binom{n}{3}^{-1} \frac{1}{3} \sum_{1\leq i_1 < i_2 < i_3 \leq n} (X_{i_3:n} - 2X_{i_2:n} + X_{i_1:n}). \tag{5.3}$$

$\hat{\lambda}_{1,n}$, is the mean of the sample. Under the conditions of Lemma 2.2, $\hat{\lambda}_{1,n}$ is a weakly consistent estimator of λ_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 λ_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 $\hat{\lambda}_{3,n}$.

$$\hat{\lambda}_{2,n} = \binom{n}{2}^{-1} \sum_{1\leq i<j\leq n} \rho(X_i, X_j), \text{ where } \rho(x, y) = |x - y|/2.$$

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

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

Theorem 5.1. *Let $\{X_n, n \ge 1\}$ 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 $\sum_{j=1}^{\infty} Cov(X_1, X_j)^{\gamma \frac{\delta}{2+\delta}} < \infty$, 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 $n \to \infty$,

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

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

$$\frac{\sqrt{n}(\hat{\lambda}_{2,n} - \lambda_2)}{2\sigma_1} \xrightarrow{\mathcal{L}} N(0, 1).$$

Also, $Cov(h^{(1)}(X_1), h^{(1)}(X_j)) \le CCov(X_1, X_j)$, $j \ge 1$.

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

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

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

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

$$\frac{\sqrt{n}(\hat{\lambda}_{3,n} - \lambda_3)}{3\sigma_2} \xrightarrow{\mathcal{L}} N(0, 1).$$

Also, $Cov(t^{(1)}(X_1), t^{(1)}(X_j)) \le CCov(X_1, X_j)^{\frac{\delta}{2+\delta}}$, $j \ge 1$.

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 $\hat{\lambda}_{r,n}$, $r = 4, \cdots, n$.

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 ($MO - LLG(\alpha, \beta, p, g)$) is,

$$\bar{F}(x; p, \alpha, \beta, g) = \frac{1}{\frac{1}{p}(\frac{x-g}{\alpha})^{\beta} + 1}, \quad \alpha, \beta, p > 0, x > g. \tag{5.4}$$

When $p = 1$, $\bar{F}$ is survival funtion of a Log-logistic ($LLG(\alpha, \beta, g)$) random variable with shape parameter β, location parameter g, and scale parameter α. 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 $MO - LLG(\alpha, \beta, p, g)$.

$$\begin{aligned} X_n &= \epsilon_n \text{ with probability } p, \\ &= min(X_{n-1}, \epsilon_n) \quad \text{with probability } (1-p),\ n \geq 1, \end{aligned} \tag{5.5}$$

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,

$$Cov(X_0, X_n) \leq (1-p)^n Var(X_0),\ \ n \geq 1. \tag{5.6}$$

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

5.3.1 Point estimation of L-moments and parameters α, β, and g

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

$$c = \frac{1}{\beta} = \frac{\lambda_3}{\lambda_2}, \quad a = \alpha p^c = \lambda_2 \frac{sin(\pi c)}{\pi c^2}, \quad g = \lambda_1 - \frac{\lambda_2^2}{\lambda_3}. \tag{5.7}$$

The conditions of Theorem 5.1 are satisfied under $\beta \geq 3$. As $n \to \infty$, $\hat{\lambda}_{r,n} \to \lambda_r$ in $\mathbb{L}_2$, $r = 1, 2, 3$. The condition $\sum_{j=1}^{\infty} Cov(X_1, X_j)^{\gamma \frac{\delta}{2+\delta}} < \infty$, 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 α, β, 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 α, β, 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 α, β, and g respectively. The following results are based on $N = 5000$ iterations.

Table 5.1 ***Estimation of*** $(\lambda_1, \lambda_2, \lambda_3, \alpha, \beta, g)$

$p = 0.3$, $\lambda_1 = 2.1680$, $\lambda_2 = 0.0336$, $\lambda_3 = 0.0067$	n=100	n=200	n=500
$\bar{\hat{\lambda}}_{1,n}(Est.M.S.E(\hat{\lambda}_{1,n}))$	2.1682(0.00010)	2.1680(0.00005)	2.1681(0.00002)
$\bar{\hat{\lambda}}_{2,n}(Est.M.S.E(\hat{\lambda}_{2,n}))$	0.0330(0.00002)	0.0333(0.00001)	0.0335(0.00000)
$\bar{\hat{\lambda}}_{3,n}(Est.M.S.E(\hat{\lambda}_{3,n}))$	0.0071(0.00001)	0.0069(0.00001)	0.0068(0.00000)
$\bar{\hat{\alpha}}_n(Est.M.S.E(\hat{\alpha}_n))$	0.2553(9.32993)	0.2086(0.01180)	0.2023(0.00141)
$\bar{\hat{\beta}}_n\Big(Est.M.S.E(\hat{\beta}_n))$	7.0618(10667.2)	5.3655(11.7929)	5.1140(1.21931)
$\bar{\hat{g}}_n\Big(Est.M.S.E(\hat{g}_n))$	1.9431(9.32904)	1.9905(0.01156)	1.9973(0.00132)
$p = 0.5$, $\lambda_1 = 2.1861$, $\lambda_2 = 0.0372$, $\lambda_3 = 0.0074$	n=100	n=200	n=500
$\bar{\hat{\lambda}}_{1,n}(Est.M.S.E(\hat{\lambda}_{1,n}))$	2.1861(0.00009)	2.1860(0.00004)	2.1861(0.00002)
$\bar{\hat{\lambda}}_{2,n}(Est.M.S.E(\hat{\lambda}_{2,n}))$	0.0370(0.00002)	0.0371(0.00001)	0.0372(0.00001)
$\bar{\hat{\lambda}}_{3,n}(Est.M.S.E(\hat{\lambda}_{3,n}))$	0.0077(0.00001)	0.0076(0.00001)	0.0075(0.00000)
$\bar{\hat{\alpha}}_n\Big(Est.M.S.E(\hat{\alpha}_n))$	0.2396(1.40422)	0.2112(0.00437)	0.2039(0.00117)
$\bar{\hat{\beta}}_n\Big(Est.M.S.E(\hat{\beta}_n))$	6.3085(1234.42)	5.3920(3.63175)	5.1356(0.93060)
$\bar{\hat{g}}_n\Big(Est.M.S.E(\hat{g}_n))$	1.9596(1.40345)	1.9883(0.00421)	1.9959(0.00111)
$p = 0.8$, $\lambda_1 = 2.2045$, $\lambda_2 = 0.04089$, $\lambda_3 = 0.0082$	n=100	n=200	n=500
$\bar{\hat{\lambda}}_{1,n}(Est.M.S.E(\hat{\lambda}_{1,n}))$	2.2046(0.00008)	2.2044(0.00004)	2.2044(0.00002)
$\bar{\hat{\lambda}}_{2,n}(Est.M.S.E(\hat{\lambda}_{2,n}))$	0.0409(0.00002)	0.04085(0.00001)	0.0409(0.00000)
$\bar{\hat{\lambda}}_{3,n}(Est.M.S.E(\hat{\lambda}_{3,n}))$	0.0082(0.00001)	0.0409(0.00000)	0.0082(0.00000)
$\bar{\hat{\alpha}}_n(Est.M.S.E(\hat{\alpha}_n))$	0.2335(0.09386)	0.2134(0.00368)	0.2054(0.00097)
$\bar{\hat{\beta}}_n(Est.M.S.E(\hat{\beta}_n))$	5.9978(77.7062)	5.4052(2.76671)	5.1613(0.72606)
$\bar{\hat{g}}_n(Est.M.S.E(\hat{g}_n))$	1.9665(0.09352)	1.9865(0.00354)	1.9946(0.00092)

In the table above, (1) $\bar{\hat{\lambda}}_{r,n} = \frac{\sum_{i=1}^{N} \hat{\lambda}_{r,n,i}}{N}$, where $\hat{\lambda}_{1,n,i}$ is estimate of λ_r in the i^{th} iteration, $i = 1, 2, \cdots, N$, $r = 1, 2, 3$.
(2) Est. M.S.E$(\hat{\lambda}_{r,n}) = \frac{\sum_{i=1}^{N} (\hat{\lambda}_{r,n,i} - \bar{\hat{\lambda}}_{r,n})^2}{N}$, $r = 1, 2, 3$.
Similarly, the values for $\bar{\hat{\alpha}}_n$, $\bar{\hat{\beta}}_n$, $\bar{\hat{g}}_n$, $Est.M.S.E(\hat{\alpha}_n)$, $Est.M.S.E(\hat{\beta}_n)$, and $Est.M.S.E(\hat{g}_n)$ were obtained.

<u>Observations</u>

(i) *Estimation of* λ_1, λ_2 *and* λ_3: For a fixed set of values of the parameters, as the sample size increases, $\bar{\hat{\lambda}}_{r,n}$, $r = 1, 2, 3$ 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) *Estimation of α, β 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 β.

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 5.2 ***Empirical probabilities (EP)***

$p = 0.3$	n=20	n=50	n= 100	n= 200
EP for the 95^{th} percentile	0.9316(0.9642)	0.9424(0.9608)	0.9464(0.9596)	0.9536(0.961)
EP for the 50^{th} percentile	0.5792(0.6568)	0.5376(0.5998)	0.5068(0.5616)	0.4986(0.5522)
EP for the 5^{th} percentile	0.2196(0.2984)	0.1624(0.2036)	0.1320(0.1606)	0.1076(0.1206)
$p = 0.5$	n=20	n=50	n= 100	n= 200
EP for the 95^{th} percentile	0.9272(0.9606)	0.9538(0.9688)	0.9568(0.9658)	0.9610(0.9654)
EP for the 50^{th} percentile	0.5394(0.6128)	0.5068(0.5748)	0.4930(0.5492)	0.4930(0.534)
EP for the 5^{th} percentile	0.1942(0.2628)	0.1456(0.1806)	0.1206(0.1406)	0.1020(0.1122)
$p = 0.8$	n=20	n=50	n= 100	n= 200
EP for the 95^{th} percentile	0.9224(0.9564)	0.9594(0.9692)	0.9664(0.9734)	0.969(0.968)
EP for the 50^{th} percentile	0.5010(0.5784)	0.4934(0.5464)	0.4724(0.5244)	0.4920(0.5258)
EP for the 5^{th} percentile	0.1776(0.2406)	0.1370(0.1636)	0.1134(0.1232)	0.1002(0.1058)

To apply Theorem 5.1, an estimator for σ_1^2 is needed. We obtained a consistent estimator by using Theorem 3.1. In general, $h^{(1)}$ would not be known. We used a consistent point estimator $\hat{h}_n^{(1)}$ of $h^{(1)}$, $\hat{h}_n^{(1)}(x) = \sum_{j=1}^{n} |X_j - x|/(2n)$. Taking $Y_i = \hat{h}_n^{(1)}(X_i)$ in (3.6), we get a consistent estimator for σ_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 5^{th} 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 5.3 ***Empirical Coverage Probabilities (CP) for*** **95%** ***CIs.***

$p = 0.3$	n=50	n=100	n=200	n=400
empirical CP for λ_1	0.8535	0.8845	0.8995	0.9220
empirical CP for λ_2	0.8345	0.8715	0.8960	0.9190
empirical CP for λ_3	0.8445	0.8820	0.9020	0.9170
$p = 0.5$	n=50	n=100	n=200	n=400
empirical CP for λ_1	0.893	0.9135	0.9325	0.9385
empirical CP for λ_2	0.8675	0.8965	0.9135	0.9250
empirical CP for λ_3	0.8200	0.8785	0.904	0.9110
$p = 0.8$	n=50	n=100	n=200	n=400
empirical CP for λ_1	0.9030	0.919	0.9265	0.9285
empirical CP for λ_2	0.878	0.9025	0.9155	0.9245
empirical CP for λ_3	0.8215	0.855	0.897	0.9100

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

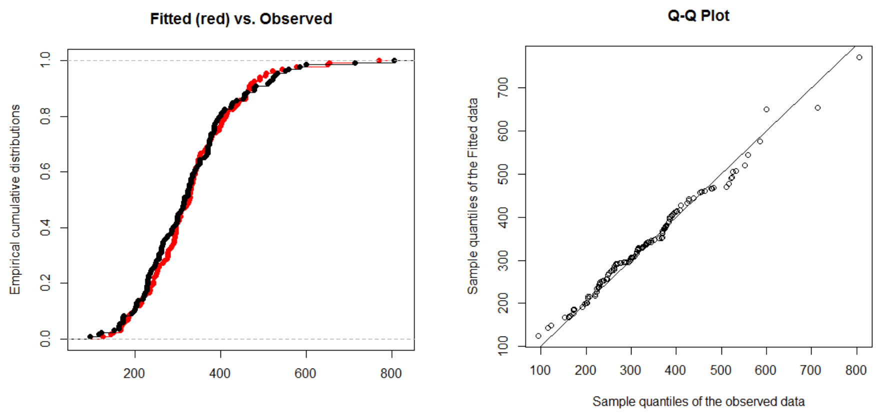
We fit the discussed MO-LLG process to the dataset consisting of the annual peak flows (in m^3/s) of The Thames at Kingston for the years $1883 - 2014$ (Source: The UK National River Flow Archive (NRFA), http://nrfa.ceh.ac.uk/data/station/peakflow/39001). The dataset consists of 132 values.

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 $\hat{\alpha} p^{1/\hat{\beta}} + \hat{g} = 313.5717$.

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

Figure 1: Fitted vs. Observed



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 λ_1: $(312.3809, 343.6756)$, for λ_2: $(54.9795, 73.7574)$, and for λ_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 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

Alice T. and Jose K. (2005). Marshall-Olkin semi-Weibull minification processes. *Recent Adv. Stat. Theory. Appl I*, (1):6–17.

Balakrishna N. and Jacob T. (2003). Parameter estimation in minification processes. *Commun. Stat. A-Theor.*, 32(11):2139–2152.

Beare B.K. (2009). A generalization of Hoeffding's lemma, and a new class of covariance inequalities. *Statist. Probab. Lett.*, 79(5):637–642.

Beutner E. and Zähle H. (2012). Deriving the asymptotic distribution of U- and V-statistics of dependent data using weighted empirical processes. *Bernoulli*, 18(3):803–822.

Beutner E. and Zähle H. (2014). Continuous mapping approach to the asymptotics of U- and V-statistics. *Bernoulli*, 20(2):846–877.

Bücher A. and Kojadinovic I. (2016). Dependent multiplier bootstraps for non-degenerate U-statistics under mixing conditions with applications. *J. Statist. Plann. Inference*, 170:83 – 105.

Bulinski A. and Shashkin A. (2007). *Limit theorems for associated random fields and related systems*. Advanced Series on Statistical Science and Applied Probability. World Scientific, Singapore.

Bulinski A. and Shashkin A. (2009). *Limit Theorems for Associated Random Variables*. Brill Academic Pub.

Dehling H. and Wendler M. (2010). Central limit theorem and the bootstrap for -statistics of strongly mixing data. *J. Multivar. Anal.*, 101(1):126 – 137.

Doukhan P., Lang G., Leucht A., and Neumann M.H. (2015). Dependent wild bootstrap for the empirical process. *J. Time Ser. Anal.*, 36(3):290–314.

Doukhan P., Prohl S., and Robert C.Y. (2011). Subsampling weakly dependent time series and application to extremes. *Test*, 20(3):447–479.

Efron B. (1979). Bootstrap Methods: Another Look at the Jackknife. *Ann. Statist.*, 7(1):1–26. doi:10.1214/aos/1176344552.

Efron B. and Tibshirani R. (1994). *An Introduction to the Bootstrap*. Chapman & Hall/CRC Monographs on Statistics & Applied Probability. Taylor & Francis.

Esary J.D., Proschan F., and Walkup D.W. (1967). Association of random variables, with applications. *Ann. Math. Statist.*, 38(5):1466–1474.

Fitzgerald D.L. (2005). Analysis of extreme rainfall using the log logistic distribution. *Stoch. Env. Res. Risk. A.*, 19(4):249–257.

Garg M. and Dewan I. (2015). On asymptotic behavior of U-statistics based on associated random variables. *Statist. Probab. Lett.*, 105:209 – 220.

Gui W. (2013). Marshall-olkin extended log-logistic distribution and its application in minification processes. *Appl. Math. Sci.*, 7(80):3947–3961.

Hoeffding W. (1948). A class of statistics with asymptotically normal distribution. *Ann. Math. Statistics*, 19(3):293–325.

Hosking J.R. (1990). L-moments: analysis and estimation of distributions using linear combinations of order statistics. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 52(1):105–124.

Hwang E. and Shin D.W. (2012). Strong consistency of the stationary bootstrap under ψ-weak dependence. *Statist. Probab. Lett.*, 82(3):488–495.

Jones M.C. and Balakrishnan N. (2002). How are moments and moments of spacings related to distribution functions? *J. Statist. Plann. Inference*, 103(1-2):377–390.

Jose K.K., Naik S.R., and Ristić M.M. (2010). Marshall-Olkin q-Weibull distribution and max–min processes. *Statist. Papers*, 51(4):837–851.

Kunsch H.R. (1989). The Jackknife and the Bootstrap for general stationary observations. *Ann. Statist.*, 17(3):1217–1241.

Lahiri S., Furukawa K., and Lee Y.D. (2007). A nonparametric plug-in rule for selecting optimal block lengths for block bootstrap methods. *Stat. Methodol.*, 4(3):292–321.

Lahiri S.N. (2003). *Resampling methods for dependent data*. Springer Science & Business Media.

Leucht A. (2012). Degenerate U- and V-statistics under weak dependence: Asymptotic theory and bootstrap consistency. *Bernoulli*, 18(2):552–585.

Lewis P.A.W. and McKenzie E. (1991). Minification processes and their Transformations. *J. Appl. Probab.*, 28(1):45–57.

Liu R.Y. and Singh K. (1992). Efficiency and Robustness in resampling. *Ann. Statist.*, 20(1):370–384.

Newman C.M. (1984). Asymptotic independence and limit theorems for positively and negatively dependent random variables. In *Inequalities in statistics and probability (Lincoln, Neb., 1982)*, volume 5 of *IMS Lecture Notes Monogr. Ser.*, pages 127–140. Inst. Math. Statist., Hayward, CA.

Oliveira P. (2012). *Asymptotics for Associated Random Variables*. Springer.

Politis D.N. and Romano J.P. (1992). A circular block-resampling procedure for stationary data. *Exploring the limits of bootstrap*, pages 263–270.

Politis D.N. and Romano J.P. (1994). The stationary bootstrap. *J. Amer. Statist. Assoc.*, 89(428):1303–1313.

Prakasa Rao B.L.S. (2012). *Associated sequences, demimartingales and nonparametric inference*. Birkhäuser/Springer Basel AG, Basel.

R Core Team (2014). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Shao J. and Tu D. (2012). *The Jackknife and Bootstrap*. Springer Science & Business Media.

Shao X. (2010). The dependent wild bootstrap. *J. Amer. Statist. Assoc.*, 105(489):218–235.

Sharipov O.S., Tewes J., and Wendler M. (2016). Bootstrap for U-statistics: a new approach. *J. Nonparametr. Statist.*, 28(3):576–594.

Sharipov O.S. and Wendler M. (2012). Bootstrap for the sample mean and for U-statistics of mixing and near-epoch dependent processes. *J. Nonparametr. Statist.*, 24(2):317–342.

Wang D., Hutson A.D., and Miecznikowski J.C. (2010). L-moment estimation for parametric survival models given censored data. *Stat Methodol.*, 7(6):655–667.

Yang T., Xu C.Y., Shao Q.X., and Chen X. (2010). Regional flood frequency and spatial patterns analysis in the Pearl River Delta region using L-moments approach. *Stoch. Env. Res. Risk. A.*, 24(2):165–182.