Stability of -statistics from weakly dependent observations

Marek Kaluszka, Andrzej Okolewski

▶ **To cite this version:**

Marek Kaluszka, Andrzej Okolewski. Stability of -statistics from weakly dependent observations. Statistics and Probability Letters, 2011, 81 (5), pp.618. 10.1016/j.spl.2010.12.015 . hal-00725106

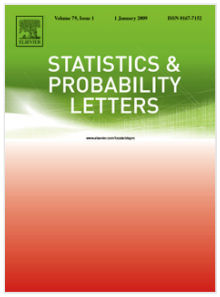
HAL Id: hal-00725106
https://hal.science/hal-00725106

Submitted on 24 Aug 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Accepted Manuscript



Stability of L-statistics from weakly dependent observations

Marek Kaluszka, Andrzej Okolewski

PII: S0167-7152(10)00361-5
DOI: 10.1016/j.spl.2010.12.015
Reference: STAPRO 5871

To appear in: *Statistics and Probability Letters*

Received date: 25 October 2010
Revised date: 20 December 2010
Accepted date: 22 December 2010

Please cite this article as: Kaluszka, M., Okolewski, A., Stability of L-statistics from weakly dependent observations. *Statistics and Probability Letters* (2011), doi:10.1016/j.spl.2010.12.015

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

Stability of L-statistics from weakly dependent observations

Marek Kaluszka[a], Andrzej Okolewski[a,*]

[a]*Institute of Mathematics, Technical University of Lodz, ul. Wolczanska 215, 90-005 Lodz, Poland*

Abstract

We study the stability of moments of L-estimates with respect to several types of weak dependence motivated by different mixing concepts. An actuarial interpretation of the presented results is indicated.

Keywords: Order statistics; L-statistics; Spectral risk measures; Weak dependence; Mixings

1. Introduction

Let $X_1, X_2, \ldots, X_n$ be non-negative random variables defined on a common probability space $(\Omega, \mathcal{F}, \mathbf{P})$. Denote by $X_{1:n} \leq X_{2:n} \leq \ldots \leq X_{n:n}$ the order statistics from the sample $X_1, \ldots, X_n$. Linear combinations of order statistics called L-statistics form an important class of estimators. A review of all the developments dealing with L-statistics is presented in Serfling (1980) and David and Nagaraja (2003). A comprehensive survey of the current knowledge about bounds for L-statistics has been made by Rychlik (1998, 2001).

In financial context L-statistics accommodate numerous indices of economic inequality as well as risk measures of actuarial science (see Davydov et al., 2007, Greselin et al., 2009). In particular, they constitute a natural class of estimators for spectral risk measures. Asymptotic properties like

*Corresponding author; tel.: (48) 42 6313859; fax.: (48) 42 6363114

Email addresses: kaluszka@p.lodz.pl (Marek Kaluszka), oko@p.lodz.pl (Andrzej Okolewski)

ACCEPTED MANUSCRIPT

strong and weak convergence of some L-statistics based on independent observations to the mentioned measures were discussed e.g. by Acerbi (2002), Dowd et al. (2008) and Greselin et al. (2009). Since the assumption of mutual independence of risks is often violated in actuarial and financial practice, it becomes desirable to introduce dependence concepts which can be relevant in these sciences. Denuit et al. (2001) proved that the sum of positively dependent risks is greater in stop-loss order than the corresponding sum of independent risks. On the other hand, Darkiewicz et al. (2005) showed that there is no strict relation between values of Wang's risk measures and such indices of dependence as Pearson's, Spearmann's and Kendall's correlation coefficients. The non-asymptotic results on L-statistics from dependent samples (e.g. Arnold and Groeneveld, 1979, Kaluszka and Okolewski, 2001, 2005, Goroncy and Rychlik, 2008) can help to recognize some properties of (closely related to each other) empirical spectral and Wang's (1996) risk measures.

In this paper we study properties of L-statistics in the case when the underlying observations are weakly dependent. We consider several types of dependence motivated by mixing concepts. In Section 2 we give some evaluations on the difference of expected L-statistics from independent and weakly dependent samples with the same univariate marginal distributions. They are expressed in terms of numerical characteristics of independent samples. In Section 3 we present similar results on stability of variances of L-statistics with respect to weak dependence. The bounds can be applied for example to robust analysis of empirical spectral risk measures as well as large claims reinsurance treaties against dependence of risks (see Remark 2).

2. Stability of expected L-estimates

Let $X_1', \ldots, X_n'$ be independent random variables such that $X_i' =^d X_i$, $i = 1, \ldots, n$, and let $X_{1:n}' \leq \ldots \leq X_{n:n}'$ be the order statistics from the sample $X_1', \ldots, X_n'$. Here and subsequently, $=^d$ means the equality in distribution. Let λ_k, $k = 1, \ldots, n$, be real numbers. Denote $L = \sum_{k=1}^n \lambda_k X_{k:n}$ and $L' = \sum_{k=1}^n \lambda_k X_{k:n}'$. We will study the stability of $\mathbf{E}\, L$ with respect to dependence structures belonging to some neighbourhoods of independence motivated by mixings (for notions of ψ and ϕ mixings see, e.g., Doukhan, 1994, p.3). They can be called weak dependencies. Let us formulate the propositions for the case of identically distributed observations, which is most frequently analysed in the literature, although the proofs provide as well the results for non-identically distributed samples. We shall assume that the integrals appearing

ACCEPTED MANUSCRIPT

in the paper exist and are finite. Moreover, we adopt the convention that $0/0 = 0$ and set

$$C_i^k = \binom{i-1}{k}\binom{n}{i}.$$

Proposition 1. *If there exists an $\varepsilon > 0$ such that*

$$\frac{|\mathbf{P}(X_{k_1} > x, \ldots, X_{k_i} > x) - \mathbf{P}(X_{k_1} > x)\ldots\mathbf{P}(X_{k_i} > x)|}{\mathbf{P}(X_{k_1} > x)\ldots\mathbf{P}(X_{k_i} > x)} \leq \varepsilon \tag{1}$$

for all $x > 0$ and $1 \leq k_1 < \ldots < k_i \leq n$, $2 \leq i \leq n$, i.e. the X_i's satisfy ψ-mixing type 1 dependence condition, then

$$|\mathbf{E}\,(L - L')| \leq \varepsilon \sum_{k=1}^{n} |\lambda_k| \sum_{i=n-k+1}^{n} C_i^{n-k} \mathbf{E}\, X'_{1:i}. \tag{2}$$

Proof. We will use the following well-known identity (cf. Blom et al., 1994, p.32)

$$\mathbf{P}\left(\sum_{i=1}^{n} \mathbf{I}_{A_i} \geq k\right) = \sum_{i=k}^{n} (-1)^{i-k} \binom{i-1}{k-1} \sum_{1 \leq k_1 < \ldots < k_i \leq n} \mathbf{P}(A_{k_1} \cap \ldots \cap A_{k_i}), \tag{3}$$

where $A_1, \ldots, A_n$ are arbitrary events and $\mathbf{I}_A$ denotes the indicator of A. Of course

$$\mathbf{P}(X_{k:n} \leq x) = 1 - \mathbf{P}\left(\sum_{i=1}^{n} \mathbf{I}(X_i > x) \geq n - k + 1\right).$$

From (3) we have

$$\mathbf{P}(X_{k:n} > x) = \sum_{i=n-k+1}^{n} (-1)^{i-(n-k+1)} \binom{i-1}{n-k} \sum_{1 \leq k_1 < \ldots < k_i \leq n} \mathbf{P}(X_{k_1} > x, \ldots, X_{k_i} > x). \tag{4}$$

Combining (4) with the analogous formula for X's leads to

$$\begin{aligned} \mathbf{E}\,(L - L') = \sum_{k=1}^{n} \lambda_k \sum_{i=n-k+1}^{n} (-1)^{i-(n-k+1)} \binom{i-1}{n-k} \int_0^{\infty} \sum_{1 \leq k_1 < \ldots < k_i \leq n} [\mathbf{P}(X_{k_1} > x, \\ \ldots, X_{k_i} > x) - \mathbf{P}(X'_{k_1} > x)\ldots\mathbf{P}(X'_{k_i} > x)]\, dx. \end{aligned} \tag{5}$$

ACCEPTED MANUSCRIPT

Hence, by (1),

$$|\mathbf{E}\,(L-L')| \le \varepsilon \sum_{k=1}^{n} |\lambda_k| \sum_{i=n-k+1}^{n} \binom{i-1}{n-k} \sum_{1\le k_1<\ldots<k_i\le n} \mathbf{E}\, X'_{1:k_1,\ldots,k_i}, \tag{6}$$

where $X'_{j:k_1,\ldots,k_i}$ denotes the jth order statistics from the sample $X'_{k_1},\ldots,X'_{k_i}$. For identically distributed observations (6) takes the form (2).

Proposition 2. *Suppose that the X_i's are ψ-mixing type 2 dependent, that is, there exists an $\varepsilon > 0$ such that*

$$\frac{|\mathbf{P}(X_{k_1}\le x,\ldots,X_{k_i}\le x) - \mathbf{P}(X_{k_1}\le x)\ldots\mathbf{P}(X_{k_i}\le x)|}{1-\mathbf{P}(X_{k_1}\le x)\ldots\mathbf{P}(X_{k_i}\le x)} \le \varepsilon \tag{7}$$

for any $x>0$ and all $1\le k_1<\ldots<k_i\le n$, $2\le i\le n$. Then

$$|\mathbf{E}\,(L-L')| \le \varepsilon \sum_{i=1}^{n} \left(\sum_{k=1}^{i} C_i^{k-1} |\lambda_k| \right) \mathbf{E}\, X'_{i:i}. \tag{8}$$

Proof. From David and Nagaraja (2003, p. 99) we have

$$\mathbf{E}\,L = \sum_{k=1}^{n} \lambda_k \int_0^\infty \left(1 - \sum_{i=k}^{n} (-1)^{i-k} \binom{i-1}{k-1} \sum_{1\le k_1<\ldots<k_i\le n} \mathbf{P}(X_{k_1}\le x,\ldots,X_{k_i}\le x) \right) dx.$$

Hence

$$\begin{aligned} \mathbf{E}\,(L-L') = \int_0^\infty \sum_{k=1}^{n} \lambda_k \sum_{i=k}^{n} (-1)^{i-k} \binom{i-1}{k-1} \sum_{1\le k_1<\ldots<k_i\le n} & [\mathbf{P}(X_{k_1}\le x)\ldots\mathbf{P}(X_{k_i}\le x) \\ - \ & \mathbf{P}(X_{k_1}\le x,\ldots,X_{k_i}\le x)]\, dx \end{aligned}$$

and, by (7),

$$\begin{aligned} |\mathbf{E}\,(L-L')| &\le \varepsilon \sum_{k=1}^{n} |\lambda_k| \sum_{i=k}^{n} \binom{i-1}{k-1} \sum_{1\le k_1<\ldots<k_i\le n} \int_0^\infty [1-\mathbf{P}(X'_{i:k_1\ldots k_i}\le x)]\, dx \\ &= \varepsilon \sum_{k=1}^{n} |\lambda_k| \sum_{i=k}^{n} \binom{i-1}{k-1} \sum_{1\le k_1<\ldots<k_i\le n} \mathbf{E}\, X'_{i:k_1,\ldots,k_i}, \end{aligned} \tag{9}$$

which gives (8).

ACCEPTED MANUSCRIPT

Proposition 3. *Assume that for an* $\varepsilon > 0$ *and all* $1 \le k_1 < \ldots < k_i \le n$, $2 \le i \le n$,

$$\sup_{x>0} \left| \mathbf{P}(X_{k_i} > x | X_{k_{i-1}} > x, \ldots, X_{k_1} > x) - \mathbf{P}(X_{k_i} > x) \right| \le \varepsilon, \tag{10}$$

i.e. the X_i*'s are* ϕ*-mixing type dependent. Then*

$$|\mathbf{E}\,(L - L')| \le \varepsilon \sum_{k=1}^{n} |\lambda_k| \sum_{i=n-k+1}^{n} C_i^{n-k} \sum_{j=2}^{i} \sum_{s=0}^{i-j} \binom{i-j}{s} \varepsilon^{i-j-s} \mathbf{E}\, X'_{1:j+s-1}. \tag{11}$$

Proof. From (5) we have

$$\begin{aligned} \mathbf{E}\,(L - L') \;=\; & \sum_{k=1}^{n} \lambda_k \sum_{i=n-k+1}^{n} (-1)^{i-(n-k+1)} \binom{i-1}{n-k} \\ & \int_0^\infty \sum_{1 \le k_1 < \ldots < k_i \le n} \left[\prod_{j=1}^{i} a_j(x) - \prod_{j=1}^{i} b_j(x) \right] dx, \end{aligned}$$

where $a_1(x) = \mathbf{P}(X_{k_1} > x)$, $a_j(x) = \mathbf{P}(X_{k_j} > x | X_{k_{j-1}} > x, \ldots, X_{k_1} > x)$ for $2 \le j \le n$ and $b_j(x) = \mathbf{P}(X'_{k_j} > x)$ for $1 \le j \le n$. Applying the identity

$$\prod_{j=1}^{i} a_j - \prod_{j=1}^{i} b_j = \sum_{j=1}^{i} (a_j - b_j) \left[\prod_{r=1}^{j-1} b_r \right] \left[\prod_{r=j+1}^{i} a_r \right]$$

with $\prod_{r=s}^{t} c_r = 1$ for $s > t$, we get

$$\begin{aligned} |\mathbf{E}\,(L - L')| \le \varepsilon \; & \sum_{k=1}^{n} |\lambda_k| \sum_{i=n-k+1}^{n} \binom{i-1}{n-k} \sum_{1 \le k_1 < \ldots < k_i \le n} \\ & \sum_{j=2}^{i} \int_0^\infty \left[\prod_{r=1}^{j-1} b_r(x) \right] \left[\prod_{r=j+1}^{i} (\varepsilon + b_r(x)) \right] dx, \end{aligned} \tag{12}$$

which for identically distributed observations leads to (11).

Next result concerns the dependence type motivated by asymptotic quadrant independence introduced by Birkel (1992).

ACCEPTED MANUSCRIPT

Proposition 4. *Let*

$$\sup_{x>0}\frac{|\mathbf{P}(X_{k_1}>x,\dots,X_{k_i}>x)-\mathbf{P}(X_{k_1}>x)\dots\mathbf{P}(X_{k_i}>x)|}{\mathbf{P}(X_{k_1}>x)\mathbf{P}(X_{k_1}\le x)\dots\mathbf{P}(X_{k_i}>x)\mathbf{P}(X_{k_i}\le x)}\le\varepsilon \tag{13}$$

for an $\varepsilon>0$ *and all* $1\le k_1<\ldots<k_i\le n$, $2\le i\le n$. *Then*

$$|\mathbf{E}\,(L-L')|\le\varepsilon\sum_{k=1}^{n}|\lambda_k|\sum_{i=k}^{n}C_i^{k-1}\mathbf{E}\left(X''_{1:i}-X'_{i:i}\right)_+, \tag{14}$$

where $x_+=\max\{0,x\}$ *and* $X'_1,\dots,X'_n$ *and* $X''_1,\dots,X''_n$ *are independent samples such that* $X'_i=^d X''_i=^d X_i$ *for* $i=1,\dots,n$.

Proof. By (5), (13) and Fubini's theorem, we obtain

$$\begin{aligned}|\mathbf{E}\,(L-L')|&\le\varepsilon\sum_{k=1}^{n}|\lambda_k|\sum_{i=n-k+1}^{n}\binom{i-1}{n-k}\int_0^\infty\sum_{1\le k_1<\ldots<k_i\le n}\\&\qquad\prod_{j=1}^{i}\mathbf{P}(X_{k_j}\le x)\mathbf{P}(X_{k_j}>x)\,dx\\&=\varepsilon\sum_{k=1}^{n}|\lambda_k|\sum_{i=n-k+1}^{n}\binom{i-1}{n-k}\sum_{1\le k_1<\ldots<k_i\le n}\\&\qquad\mathbf{E}\int_0^\infty\mathbf{I}(X'_{i:k_1,\dots,k_i}\le x)\mathbf{I}(X''_{1:k_1,\dots,k_i}>x)\,dx\\&=\varepsilon\sum_{k=1}^{n}|\lambda_k|\sum_{i=k}^{n}\binom{i-1}{k-1}\sum_{1\le k_1<\ldots<k_i\le n}\mathbf{E}\left(X''_{1:k_1,\dots,k_i}-X'_{i:k_1,\dots,k_i}\right)_+\end{aligned} \tag{15}$$

and this implies (14).

Example 1. Let $X_1,\dots,X_n$ have the standard exponential distribution. Then $\mathbf{E}\,X'_{k:n}=\sum_{i=n-k+1}^{n}1/i$ (see David and Nagaraja, 2003, p. 52), and consequently the bounds of Propositions 1,2,3 and 4, respectively, take the forms

$$|\mathbf{E}\,(X_{k:n}-X'_{k:n})|\le\varepsilon\sum_{i=n-k+1}^{n}C_i^{n-k}\frac{1}{i},$$

$$|\mathbf{E}\,(X_{k:n}-X'_{k:n})|\le\varepsilon\sum_{i=k}^{n}C_i^{k-1}\sum_{j=1}^{i}\frac{1}{j},$$

ACCEPTED MANUSCRIPT

$$|\mathbf{E}\,(X_{k:n} - X'_{k:n})| \le \varepsilon \sum_{i=n-k+1}^{n} C_i^{n-k} \sum_{j=2}^{i} \sum_{s=0}^{i-j} \binom{i-j}{s} \frac{\varepsilon^{i-j-s}}{j+s-1}$$

and

$$|\mathbf{E}\,(X_{k:n} - X'_{k:n})| \le \varepsilon \sum_{i=k}^{n} C_i^{k-1} \left[i \binom{2i}{i} \right]^{-1}.$$

Remark 1. The formulae (6), (9), (12) and (15) provide the bounds for the case when the weakly dependent observations $X_1, \ldots, X_n$ are non-identically distributed.

Dependence neighbourhoods (1), (10) and (13) for $n = 2$ contain some important classes of copulas. For example, the Farlie-Gumbel-Morgenstern, Ali-Mikhail-Haq and Plackett families of copulas (see Matuła, 2005) with parameters from the intervals $[0, \varepsilon]$, $[0, \varepsilon/(1+\varepsilon)]$ and $[1, 1+\varepsilon)$, respectively, belong to these neighbourhoods.

Instead of using Kolmogorov-type neighbourhoods of independence, one can consider the other types of neighbourhoods. Let us give one such example for the case of $n = 2$, which is intensively studied in reliability theory (see e.g. Balakrishnan and Lai, 2009). We have

$$\mathbf{E}\,(L - L') = (\lambda_1 - \lambda_2)(\mathbf{E}\,X_{1:2} - \mathbf{E}\,X'_{1:2}) = (\lambda_2 - \lambda_1)(\mathbf{E}\,X_{2:2} - \mathbf{E}\,X'_{2:2}),$$

since $X_{2:2} = X_1 + X_2 - X_{1:2}$. Obviously,

$$\mathbf{E}\,X_{2:2} - \mathbf{E}\,X'_{2:2} = \int_0^\infty \Big[\mathbf{P}(X'_{2:2} \le t) - \mathbf{P}(X_{2:2} \le t) \Big]\, dt. \tag{16}$$

If $F_1 = F_2 = F$, then

$$|\mathbf{E}\,(L - L')| \le |\lambda_2 - \lambda_1| \int_0^1 |C(t,t) - t^2|\, dt,$$

where $C(t,t) = \mathbf{P}(X_1 \le F^{-1}(t), X_2 \le F^{-1}(t))$. In the case when C is a generalized Farlie-Gumbel-Morgenstern copula, i.e.

$$C(x,y) = xy + \rho x^b y^b (1-x)^a (1-y)^a, \quad a, b \ge 1,\ |\rho| \le 1,$$

we obtain

$$|\mathbf{E}\,(L - L')| \le |\rho||\lambda_2 - \lambda_1| B(2a+1, 2b+1),$$

where $B(x,y)$ denotes the Euler beta function. In particular, if $|\rho| \le \varepsilon$ for some ε, then $|\mathbf{E}\,(L - L')| \le \varepsilon|\lambda_2 - \lambda_1|/3$ for arbitrary $a, b \ge 1$, since $\int_0^1 x^{2a}(1-x)^{2b}\, dx \le \min\{1/(2a+1), 1/(2b+1)\}$.

ACCEPTED MANUSCRIPT

Remark 2. The presented inequalities can be applied to evaluate the weak dependent insurance risks. Suppose that $X_1, \ldots, X_n$ is a homogeneous portfolio of risks satisfying ψ-mixing condition (7). If an insurer is able to determine the parameter ε for this mixing type dependence model and if he is willing to use some empirical spectral risk measure $\sum_{i=1}^{n} \lambda_i X_{i:n}$ with some $0 \leq \lambda_1 \leq \ldots \leq \lambda_n$, then the bound (8) says that the additional (in comparison to the case of independent risks) loading related to the dependence of risks cannot exceed the coefficient determined by independent risks multiplied by ε. If the dependence is really weak, i.e. ε is close to zero, then the loading should be small. Another application of L-statistics to insurance is the large claim reinsurance treaty (cf. Kremer, 1998). For this reinsurance contract not only the net premium but also the risk measured by variance may be of interest.

3. Stability of second moments of *L*-statistics

We will restrict our attention to one mixing type of dependence. The other mixings may be treated analogously. Let us examine first the variances of single order statistics.

Proposition 5. *Assume that there exists an $\varepsilon > 0$ such that inequality (1) is satisfied for all $x > 0$ and all $1 \leq k_1 < \ldots < k_i \leq n$, $2 \leq i \leq n$. Then for any k*

$$\left|\mathbf{Var}X_{k:n} - \mathbf{Var}X'_{k:n}\right| \leq \varepsilon \sum_{i=n-k+1}^{n} C_i^{n-k} \left(\mathbf{E}\,[X'_{1:i}]^2 + 2\mathbf{E}\,X'_{1:i}\mathbf{E}\,X'_{k:n}\right) + \varepsilon^2 \left[\sum_{i=n-k+1}^{n} C_i^{n-k}\mathbf{E}\,X'_{1:i}\right]^2. \tag{17}$$

Proof. From Hoeffding's identity (see, e.g., Szekli, 1995, p. 136) for non-negative random variables

$$\mathbf{Var}X = \int_{\mathbf{R}_+^2} \left[\mathbf{P}(X > x \vee y) - \mathbf{P}(X > x)\mathbf{P}(X > y)\right]\, dx\, dy,$$

we obtain

$$\mathbf{Var}X_{k:n} - \mathbf{Var}X'_{k:n} = \int_{\mathbf{R}_+^2} \left[\mathbf{P}(X_{k:n} > x \vee y) - \mathbf{P}(X'_{k:n} > x \vee y)\right]\, dx\, dy$$

ACCEPTED MANUSCRIPT

$$+ \int_{\mathbf{R}_+^2} \mathbf{P}(X'_{k:n} > y) \left[\mathbf{P}(X'_{k:n} > x) - \mathbf{P}(X_{k:n} > x)\right] \, dx\, dy$$
$$+ \int_{\mathbf{R}_+^2} \left[\mathbf{P}(X_{k:n} > x) - \mathbf{P}(X'_{k:n} > x)\right] \left[\mathbf{P}(X'_{k:n} > y) - \mathbf{P}(X_{k:n} > y)\right] \, dx\, dy$$
$$+ \int_{\mathbf{R}_+^2} \mathbf{P}(X'_{k:n} > x) \left[\mathbf{P}(X'_{k:n} > y) - \mathbf{P}(X_{k:n} > y)\right] \, dx\, dy, \tag{18}$$

where $x \vee y = \max\{x, y\}$. By (4) and (1),

$$\left| \int_0^\infty [\mathbf{P}(X_{k:n} > y) - \mathbf{P}(X'_{k:n} > y)] \, dy \right| \leq \varepsilon \sum_{i=n-k+1}^{n} \binom{i-1}{n-k} \sum_{1 \leq k_1 < \ldots < k_i \leq n} \int_0^\infty \mathbf{P}(X'_{k_1,\ldots,k_i} > y) \, dy. \tag{19}$$

Combining (18) and (19) we get

$$\begin{aligned} \left| \mathbf{Var} X_{k:n} - \mathbf{Var} X'_{k:n} \right| \;\leq\; & 2\varepsilon \sum_{i=n-k+1}^{n} \binom{i-1}{n-k} \sum_{1 \leq k_1 < \ldots < k_i \leq n} \int_0^\infty \int_y^\infty \mathbf{P}(X'_{1:k_1,\ldots,k_i} > x) \, dx\, dy \\ & + 2\varepsilon \mathbf{E}\, X'_{k:n} \sum_{i=n-k+1}^{n} \binom{i-1}{n-k} \sum_{1 \leq k_1 < \ldots < k_i \leq n} \mathbf{E}\, X'_{1:k_1,\ldots,k_i} \\ & + \varepsilon^2 \left[\sum_{i=n-k+1}^{n} \binom{i-1}{n-k} \sum_{1 \leq k_1 < \ldots < k_i \leq n} \mathbf{E}\, X'_{1:k_1,\ldots,k_i} \right]^2. \end{aligned}$$

Applying the identity: $2 \int_0^\infty \int_y^\infty \mathbf{P}(X > x) \, dx\, dy = \mathbf{E}\, X^2$, leads to (17).

Under a stronger mixing condition, the result of Proposition 5 can be extended to L-statistics.

Proposition 6. *Suppose that there is an $\varepsilon > 0$ such that*

$$\sup_{x,y>0} \left| \frac{\mathbf{P}(X_{k_1} > x_1, \ldots, X_{k_i} > x_i)}{\mathbf{P}(X_{k_1} > x_1) \ldots \mathbf{P}(X_{k_i} > x_i)} - 1 \right| \leq \varepsilon, \tag{20}$$

where $x_j = x$ or $x_j = y$, $j = 1, \ldots, i$, $1 \leq k_1 < \ldots < k_i \leq n$ and $2 \leq i \leq n$. Then

(i) *for $k < l$*

$$\left| \mathbf{Cov}(X_{k:n}, X_{l:n}) - \mathbf{Cov}(X'_{k:n}, X'_{l:n}) \right| \leq \varepsilon B_{kl},$$

ACCEPTED MANUSCRIPT

in which

$$B_{kl} = w_l \mathbf{E}\, X'_{k:n} + w_k \mathbf{E}\, X'_{l:n} + \frac{1}{2}\,(v_k + v_{kl}) + \varepsilon w_k w_l, \tag{21}$$

where w_k, v_k and v_{kl} are defined below by (24), (27) and (31), respectively;
(ii)

$$|\mathbf{Var}L - \mathbf{Var}L'| \le \varepsilon \left[\sum_{k=1}^{n} \lambda_k^2 B_{kk} + 2 \sum_{k=1}^{n-1} \sum_{l=k+1}^{n} |\lambda_k \lambda_l| B_{kl} \right] \tag{22}$$

with $B_{kk} = 2w_k \mathbf{E}\, X'_{k:n} + v_k + \varepsilon w_k^2$.

Proof. By Hoeffding's identity,

$$\mathbf{Cov}(X_{k:n}, X_{l:n}) - \mathbf{Cov}(X'_{k:n}, X'_{l:n}) = I_1 + I_2 + I_3 + I_4,$$

where

$$\begin{aligned}
I_1 &= \int_{\mathbf{R}_+^2} \Big[\mathbf{P}(X_{k:n} > x, X_{l:n} > y) - \mathbf{P}(X'_{k:n} > x, X'_{l:n} > y) \Big]\, dx\, dy, \\
I_2 &= \int_{\mathbf{R}_+^2} \mathbf{P}(X'_{k:n} > x) \Big[\mathbf{P}(X'_{l:n} > y) - \mathbf{P}(X_{l:n} > y) \Big]\, dx\, dy, \\
I_3 &= \int_{\mathbf{R}_+^2} \mathbf{P}(X'_{l:n} > y) \Big[\mathbf{P}(X'_{k:n} > x) - \mathbf{P}(X_{k:n} > x) \Big]\, dx\, dy, \\
I_4 &= \int_{\mathbf{R}_+^2} \Big[\mathbf{P}(X_{l:n} > y) - \mathbf{P}(X'_{l:n} > y) \Big] \Big[\mathbf{P}(X'_{k:n} > x) - \mathbf{P}(X_{k:n} > x) \Big]\, dx\, dy.
\end{aligned}$$

Let us evaluate each I_j, $j = 1, \ldots, 4$, separately. Proceeding analogously as in the proof of Proposition 5 we obtain

$$|I_2| \le \varepsilon \mathbf{E}\, X'_{k:n} w_l, \qquad |I_3| \le \varepsilon\, \mathbf{E}\, X'_{l:n} w_k \qquad \textit{and} \qquad |I_4| \le \varepsilon^2 w_k w_l, \tag{23}$$

in which

$$w_j = \sum_{i=n-j+1}^{n} \binom{i-1}{n-j} \sum_{1 \le k_1 < \ldots < k_i \le n} \mathbf{E}\, X'_{1:k_1,\ldots,k_i}. \tag{24}$$

ACCEPTED MANUSCRIPT

Observe that

$$I_1 = \int_0^\infty \int_0^y \Big[\mathbf{P}(X_{k:n} > x, X_{l:n} > y) - \mathbf{P}(X'_{k:n} > x, X'_{l:n} > y)\Big]\, dx\, dy + \int_0^\infty \int_y^\infty \Big[\mathbf{P}(X_{k:n} > x) - \mathbf{P}(X'_{k:n} > x)\Big]\, dx\, dy \equiv I_{11} + I_{12}. \quad (25)$$

Applying formula (4) to the second integral of (25) gives

$$|I_{12}| \leq \varepsilon \sum_{i=n-k+1}^{n} \binom{i-1}{n-k} \sum_{1\leq k_1<\ldots<k_i\leq n} \int_0^\infty \int_y^\infty \mathbf{P}(X'_{k_1} > x, \ldots, X'_{k_i} > x)\, dxdy = \varepsilon v_k/2, \quad (26)$$

where

$$v_k = \sum_{i=n-k+1}^{n} \binom{i-1}{n-k} \sum_{1\leq k_1<\ldots<k_i\leq n} \mathbf{E}\left[X'_{1:k_1,\ldots,k_i}\right]^2. \quad (27)$$

In order to evaluate the first integral I_{11} of (25) we use the identity

$$\mathbf{I}(x_{k:n} > x) = \sum_{i=n-k+1}^{n} (-1)^{i-(n-k+1)} \binom{i-1}{n-k} \sum_{1\leq k_1<\ldots<k_i\leq n} \mathbf{I}(x_{k_1} > x, \ldots, x_{k_i} > x),$$

which can be deduced from (4) by putting $X_k = x_k$ for $k = 1, \ldots, n$. It follows that

$$\mathbf{P}(X_{k:n} > x, X_{l:n} > y) = \mathbf{E} \sum_{i=n-k+1}^{n} \sum_{j=n-l+1}^{n} a_{ij} \sum_{1\leq k_1<\ldots<k_i\leq n} \sum_{1\leq l_1<\ldots<l_j\leq n} \mathbf{I}(X_{k_1} > x, \ldots, X_{k_i} > x, X_{l_1} > y, \ldots, X_{l_j} > y)$$

with

$$a_{ij} = (-1)^{i+j-(n-k+1)-(n-l+1)} \binom{i-1}{n-k} \binom{j-1}{n-l}.$$

If $k_s = l_z$ for some s, z, then $\mathbf{I}(X_{k_s} > x, X_{l_z} > y) = \mathbf{I}(X_{k_s} > y)$ for $x < y$ (cf. the definition of I_{11} in (25)). Therefore

$$|\mathbf{P}(X_{k:n} > x, X_{l:n} > y) - \mathbf{P}(X'_{k:n} > x, X'_{l:n} > y)|$$

ACCEPTED MANUSCRIPT

$$\leq \varepsilon \sum_{i=n-k+1}^{n} \sum_{j=n-l+1}^{n} |a_{ij}| \sum_{1\leq k_1<\ldots<k_i\leq n} \sum_{1\leq l_1<\ldots<l_j\leq n} \mathbf{P}\left(\bigcap_{s\in S_{ij}} \{X_s^{'} > x\} \cap \bigcap_{z\in Z_j} \{X_z^{'} > y\} \right), \tag{28}$$

where $Z_j = \{l_1, \ldots, l_j\}$ and $S_{ij} = \{k_1, \ldots, k_i\} \setminus Z_j$. Since

$$\begin{aligned}
&\int_0^{\infty}\int_0^{y} \mathbf{P} \left(\bigcap_{s\in S_{ij}} \{X_s^{'} > x\} \cap \bigcap_{z\in Z_j} \{X_z^{'} > y\} \right) dx\, dy \\
&= \mathbf{E} \int_0^{\infty} \mathbf{I}(X_{1:Z_j}^{'} > y) \int_0^{y} \mathbf{I}(X_{1:S_{ij}}^{'} > x)\, dx\, dy \\
&= \mathbf{E} \int_0^{\infty} \mathbf{I}(X_{1:Z_j}^{'} > y) \min\left(y, X_{1:S_{ij}}^{'}\right)\, dy \\
&= \int_0^{\infty} \mathbf{P}(X_{1:Z_j}^{'} > y) \mathbf{E} \min\left(y, X_{1:S_{ij}}^{'}\right)\, dy \\
&\leq \int_0^{\infty} y \mathbf{P}(X_{1:Z_j}^{'} > y)\, dy = \frac{1}{2}\mathbf{E}\left[X_{1:Z_j}^{'}\right]^2 \leq \frac{1}{2}\mathbf{E}\left[X_{n:n}^{'}\right]^2 < \infty,
\end{aligned} \tag{29}$$

we have, by (25), (28) and (29), that

$$|I_{11}| \leq \varepsilon v_{kl}/2, \tag{30}$$

with

$$v_{kl} = \sum_{i=n-k+1}^{n} \binom{n}{i} \sum_{j=n-l+1}^{n} |a_{ij}| \sum_{1\leq l_1<\ldots<l_j\leq n} \mathbf{E}\left[X_{1:Z_j}^{'}\right]^2, \tag{31}$$

where $X_{1:A}^{'}$ is the minimal order statistics from the sample of all the $X_i^{'}$s which indices are in the set A. Combining (30) with (26), (25) and (23) leads to (i). The part (ii) is a straightforward consequence of (i) and the last formula in the proof of Proposition 5.

Example 2. If $X_1, \ldots, X_n$ have the standard exponential distribution, then the coefficients in the definitions of B_{kk} and B_{kl} in Proposition 6 take the forms $w_k = \sum_{i=n-k+1}^{n} C_i^{n-k}/i$, $v_k = 2\sum_{i=n-k+1}^{n} C_i^{n-k}/i^2$ and $v_{kl} = 2\sum_{i=n-k+1}^{n} \sum_{j=n-l+1}^{n} C_i^{n-k} C_j^{n-l}/j^2$.

ACCEPTED MANUSCRIPT

Example 3. Let B denote the right-hand side of (22). Combining the Chebychev inequality with (22) we obtain the following confidence interval $[L - ((\mathbf{Var}L' + B)/\alpha)^{1/2}, L + ((\mathbf{Var}L' + B)/\alpha)^{1/2}]$ for the expected value of L-statistic from dependent observations with confidence coefficient greater than or equal to $1 - \alpha$.

Acknowledgements: The authors would like to thank the referees for their valuable comments which led to improvements in the paper. The research was supported by the Polish Ministry of Science and Higher Education Grant no. N N111 431337.

References

[1] Acerbi, C. (2002) Spectral measures of risk: A coherent representation of subjective risk aversion. *J. Banking Finance 26*, 1505–1518.

[2] Arnold, B.C., Groeneveld, R.A. (1979) Bounds on expectations of linear systematic statistics based on dependent samples. *Ann. Statist. 7*, 220–223.

[3] Balakrishnan, N., Lai, C.D. (2009) *Continuous Bivariate Distributions*. 2ed. Springer

[4] Birkel, T. (1992) Laws of large numbers under dependence assumptions. *Statist. Probab. Lett. 7*, 17–19.

[5] Blom, G., Holst, L., Sandell, D. (1994) *Problems and Snapshots from the World of Probability*. Springer, New York.

[6] Darkiewicz, G., Dhane, J., Goovaerts, M. (2005) Risk measures and dependencies of risks. *Braz. J. Probab. Stat. 19*, 155–178.

[7] David, H. A., Nagaraja, H. N. (2003) *Order Statistics*. Wiley, Hoboken.

[8] Davydov, Y., Khoshnevisian, D., Shi, Z., Zitikis, R. (2007) Convex rearrangements, generalized Lorenz curves, and Gaussian data. *J. Statist. Plann. Inference 137*, 915–934.

[9] Denuit, M. Dhane, J., Ribas, C. (2001) Does positive dependence between individual risks increase stop-loss premiums? *Insurance Math. Econom. 28*, 305–308.

ACCEPTED MANUSCRIPT

[10] Doukhan, P. (1994) *Mixing: Properties and Examples.* Springer, New York.

[11] Dowd, K., Cotter, J., Sorwar, G. (2008) Spectral Risk Measures: Properties and Limitations. *J. Finan. Serv. Res. 34*, 61–75.

[12] Goroncy, A., Rychlik, T. (2008) Lower bounds on expectations of positive L-statistics from without-replacement models. *J. Statist. Plann. Inference 138*, 3647–3659.

[13] Greselin, F., Puri, M. L., Zitikis, R. (2009) L-functions, processes, and statistics in measuring economic inequality and actuarial risk. *Statist. Interface 2*, 227–245.

[14] Kaluszka, M., Okolewski, A. (2001) An extension of the Erdős–Neveu–Rényi theorem with applications to order statistics. *Statist. Probab. Lett. 52*, 181–186.

[15] Kaluszka, M., Okolewski, A. (2005) Bounds for L-statistics from weakly dependent samples of random length. *Comm. Statist. Theory Methods 34*, 1899–1910.

[16] Kremer, E. (1998) Largest claims reinsurance premiums under possible claims dependence. *ASTIN Bulletin 28*, 257–267.

[17] Matuła, P. (2005) On some families of AQSI random variables and related strong law of large numbers. *Applied Mathematics E-Notes 5*, 31–35.

[18] Rychlik, T. (1998) Bounds on expectations of L-estimates, in: *Order statistics: Theory and Methods*, (N. Balakrishnan and C.R. Rao, eds.), Handbook of Statistics, Vol 16, North-Holland, Amsterdam, 105-145.

[19] Rychlik, T. (2001) *Projecting Statistical Functionals* Springer, New York.

[20] Serfling, R. J. (1980) *Approximation Theorems of Mathematical Statistics.* Wiley, New York.

[21] Szekli, R. (1995) *S*tochastic Ordering and Dependence in Applied Probability. Springer, New York.

ACCEPTED MANUSCRIPT

[22] Wang, S. (1996) Premium calculation by transforming the layer premium density. *Astin Bull. 16*, 71–92.

ACCEPTED MANUSCRIPT