SHARP PROBABILITY INEQUALITIES AND CONSERVATIVE TESTING PROCEDURES FOR STUDENTIZED PROCESSES AND MOVING AVERAGES WITH APPLICATIONS TO ECONOMETRIC MODELS AND HEAVY TAILS[1]

By Victor H. de la Peña and Rustam Ibragimov [2]

Department of Statistics, Columbia University, New York, NY 10027

Department of Economics, Yale University, New Haven, CT 06511

Abstract

In this paper we present an approach for developing conservative testing procedures for dependent and/or heavy tailed observations. In particular, we obtain conservative critical values for tests of independence, nonparametric $t-$tests and permutation tests against serial correlation. Power and size properties of the tests are illustrated through a Monte-Carlo study. The results are based on sharp extensions of probability and moment inequalities for sums of independent random variables to the case of moving averages and their self-normalized and Studentized analogues in independent symmetric variables. The case of statistics in dependent variables is treated through the use of measures of dependence. The results are developed on the basis of general decoupling principles of independent interest.

JEL Classification: C13, C14

Key words and phrases: Self-normalized processes; Studentized statistics; Random polynomials; Moving averages; Sample cross-moments; Sums of multilinear forms; Probability inequalities; Exponential inequalities; Moment inequalities; Time series; Serial dependence; $t-$tests; Conservative tests.

0. Introduction

A number of problems in nonparametric inference in statistics and econometrics involve estimating the tail probabilities of test statistics. Several studies have discussed applications of semiparametric and nonparametric bounds for the $p-$values of test statistics in several contexts including nonparametric $t-$tests, Hotelling's T^2 test, sign tests and signed rank tests, permutation tests against serial dependence (see, e.g., Efron (1969), Eaton and Efron (1970), Edelman (1986, 1980), Dufour and Hallin (1991, 1992, 1993) and Pinelis (1994)).

The interest in estimates for the tail probabilities of commonly used test statistics is motivated in part by the fact that the exact distributions of the statistics are frequently unknown. Even if known, the exact distributions of the test statistics are usually difficult to compute and have to be obtained by relying on computationally intensive algorithms or Monte-Carlo techniques, as in the case of permutation $t-$tests or

[1]The authors are grateful to the participants at the Prospectus Workshop at the Department of Economics, Yale University for helpful comments. V. H. de la Peña gratefully acknowledges the partial support under the NSF grants DMS/99/72237 and DMS/02/05791. R. Ibragimov gratefully acknowledges the financial support from the Yale University Graduate Fellowship and the Cowles Foundation Prize.

[2]Corresponding author. Rustam Ibragimov, 28 Hillhouse Ave., New Haven, CT 06511; Phone: (203) 777-2750; Fax: (203) 432-2128; Email: rustam.ibragimov@yale.edu

linear signed rank statistics (see Dufour and Hallin (1992, 1993)). Furthermore, large sample approximations, e.g., normal approximations, require special regularity assumptions on the distribution of the observations such as existence of the second or higher moments or identical distribution.

One should also note the importance of estimates for the tail probabilities of statistics (that hold under minimal assumptions) in the context of statistical inference in models driven by innovations with heavy tailed distributions. The latter problems are closely related to the study of robustness of tests to non-normality and fat-tailedness assumptions (see, e.g., Efron (1969), Kariya and Sinha (1988) and Jurečkovà and Sen (1996)). As it is well documented in the empirical finance literature, many financial market time series have fat tailed distributions and do not satisfy moment assumptions required for normal convergence (see, e.g., Loretan and Phillips (1994) and references therein). In studies of the latter models it is usually assumed that the error distributions belong to the domain of attraction of stable laws. The limiting distributions for test statistics in such setups are non-standard and involve functionals of stable processes, therefore, one has to rely on computationally intensive Monte-Carlo simulations to compute the critical values of the tests. Furthermore, the convergence of the test statistics in such models is very slow, hence providing inadequate approximations for finite samples (see, e.g., Adler, Feldman and Gallagher (1998) and Tse and Zhang (2000)).

In this paper, we present an approach to developing conservative testing procedures based on moving averages in independent variables and their self-normalized and Studentized counterparts. We also provide extensions to the case of dependent random variables through the use of measures of dependence. The results are of particular importance in situations when the observations exhibit heavy tails. In particular, we show how to obtain conservative critical values for: tests for independence, nonparametric $t-$tests and permutation tests against serial correlation. Power and size properties of the tests are illustrated through a Monte-Carlo study. The results are based on sharp extensions of several (of the best known) probability and moment inequalities for sums of independent symmetric random variables to the case of moving averages and their self-normalized and Studentized analogues. These extensions are developed on the basis of general decoupling principles of independent interest that can also be used for the development of asymptotic results. Finally, we cite recent results for self-normalized sums of dependent variables in de la Peña (1999) and de la Peña, Klass and Lai (2000, 2002) which provide applications to self-normalized LIL's.

Throughout the paper we focus on two general structures. The first concerns moving averages (and self-normalized counterparts under symmetry assumptions) in independent r.v.'s and the second involves extensions to the case of statistics in dependent r.v.'s using measures of dependence. For ease of reference, the paper is organized as follows: Section 1 presents a survey of probability inequalities for sums of independent r.v's. Section 2 contains the main inequalities obtained in the paper with Section 2a dealing with sharp probability inequalities for moving averages and their self-normalized and Studentized versions in independent r.v.'s and Section 2b dealing with extensions of the results in Sections 1 and 2a to the case of dependent r.v.'s through measures of dependence. Section 2c contains sharp moment inequalities for moving averages and their applications. In Section 3 and 4 we present applications of our results in testing procedures. Section 3a contains applications in permutation tests against serial correlation and in tests for independence. Section 3b contains a Monte-Carlo study of the size and power of tests considered in Section 3a. Section 4 deals with applications in testing procedures for dependent r.v.'s. In particular, in Section 4a we consider conservative critical regions for nonparametric $t-$tests and Section 4b deals with applications in conservative tests of linear hypotheses. Finally, Section 5 is devoted to proofs. In order to facilitate the reading we have made the applications section (Section 3) self-contained.

1. Sharp probability inequalities for sums of independent r.v.'s

Let $X_1, \ldots, X_n$ be random variables (r.v.'s) on a probability space $(\Omega, \Im, P)$. A question of key interest in the calculation of p-values is to accurately estimate the tail probabilities $P(\sum_{i=1}^n X_i > x)$, $x \in \mathbf{R}$. There are several results approximating tail probabilities. As examples we cite the works of Bernstein, Prokhorov, Bennett, Hoeffding and Eaton (see Bernstein (1946), Prokhorov (1959), Bennett (1962), Hoeffding (1963), Eaton (1970, 1974), Edelman (1986) and Talagrand (1996)). In what follows we present a review of the inequalities that we will be citing as well as the results for which we will provide extensions in later sections.

1. Hoeffding's inequalities (Hoeffding (1963), see also Edelman (1986)). Let $X_1, \ldots, X_n$ be independent r.v.'s with $EX_i = 0$, $i = 1, \ldots, n$, such that $|X_i| \le d_i \in \mathbf{R}$ (a.s.), $i = 1, \ldots, n$, and let $D^2 = \sum_{i=1}^n d_i^2$. Then

$$P(\sum_{i=1}^n X_i > x) \le exp(-\frac{x^2}{2D^2}), \tag{1}$$

$x > 0$.

Let Z be the standard normal r.v., $\phi(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}$, $\Phi(u) = \int_{-\infty}^u \phi(t)dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^u e^{-\frac{t^2}{2}} dt$, and let K be the class of twice differentiable even functions $f : \mathbf{R} \to \mathbf{R}$ such that f'' is nonnegative and convex and $\overline{K}$ be the class of functions $f \in K$ such that $f : \mathbf{R}_+ \to \mathbf{R}$ is a nondecreasing function. The classes K and $\overline{K}$ are quite wide and contain, for example, the functions $f(x) = |x|^t$, $t \ge 3$; $f(x) = (|x| - u)_+^t$, $t \ge 3$, $u \ge 0$ (here and in what follows, $w_+ = max(w, 0)$, $w \in \mathbf{R}$); $f(x) = e^{h|x|}$, $h > 0$, and $f(x) = cosh\, hx$, $h \ne 0$.

2. Eaton-type inequalities (implied by the estimates in Eaton (1970, 1974), see also Pinelis (1994)). Let $X_1, \ldots, X_n$ be independent r.v.'s with $EX_i = 0$, $i = 1, \ldots, n$, such that $|X_i| \le d_i \in \mathbf{R}$ (a.s.), $i = 1, \ldots, n$, and let $D^2 = \sum_{i=1}^n d_i^2$. Then the following inequalities hold:

$$P(\sum_{i=1}^n X_i > x) \le \frac{1}{2} \frac{Ef(|Z|)}{f(\frac{x}{D})}, \tag{2}$$

$f \in \overline{K}$, $x > 0$,

$$P(\sum_{i=1}^n X_i > x) \le \inf_{0<u<x/D} \int_u^\infty ((t-u)^3/(\frac{x}{D} - u)^3)\phi(t)dt, \tag{3}$$

$x > 0$.

The use of the results in Edelman (1990) and Hunt (1955) gives that for all independent mean-zero r.v.'s $X_1, \ldots, X_n$ such that $|X_i| \le d_i \in \mathbf{R}$ (a.s.), $i = 1, \ldots, n$,

$$P(\sum_{i=1}^{n} X_i > x) \le 1 - \Phi(\frac{x}{D} - \frac{1.5D}{x}), \tag{4}$$

$x > 0$ (see the proof of Theorem 1 in this paper). Pinelis (1994) obtained the following estimates:

$$P(\sum_{i=1}^{n} X_i > x) \le \frac{2e^3}{9}(1 - \Phi(\frac{x}{D})) \le \frac{e^3}{9}\frac{\phi(\frac{x}{D})D}{x}, \tag{5}$$

$x > 0$ (the second inequality in (5) was conjectured by Eaton (1974)). Pinelis (1994) also proposed the following alternative to (3):

$$P(\sum_{i=1}^{n} X_i > x) \le min(1/2, D^2/(2x^2), \inf_{0<u<x/D} \int_u^{\infty} ((t-u)^3/(\frac{x}{D} - u)^3)\phi(t)dt), \tag{6}$$

Dufour and Hallin (1993) noted that bounds (3) and (6) can be improved when the number of the r.v.'s is taken into account and proved an inequality from which it follows that under the above conditions,

$$P(\sum_{i=1}^{n} X_i > x) \le min\left(1/2, D^2/(2x^2), B(x/D, n)\right), \tag{7}$$

where

$$B(y, n) = 2^{1-n} \inf_{0 \le c < y} \sum_{m=0}^{n} C_n^m f_c[(n/4)^{-1/2}(m - (n/2))]/(y - c)^3, \tag{8}$$

$f_c(t) = [(|t| - c)_+]^3$, $C_n^m = n!/(m!(n-m)!)$. Inequalities (2)-(7), with the right-hand side expressions multiplied by 2, hold for $|\sum_{i=1}^{n} X_i|$ as well.

From the results obtained by Eaton (1970, 1974) it also follows that the following inequality holds:

$$Ef(\sum_{i=1}^{n} c_i X_i) \le Ef(Z) \tag{9}$$

for all $f \in K$, independent r.v.'s $X_1, \ldots, X_n$ with $EX_i = 0$, $i = 1, \ldots, n$, such that $|X_i| \le 1$ (a.s.), $i = 1, \ldots, n$, and constants $c_i \in \mathbf{R}$, $i = 1, \ldots, n$, such that $\sum_{i=1}^{n} c_i^2 = 1$ (see also Pinelis (1994)).

Edelman (1986, 1990) and Pinelis (1994) applied inequalities (1), (3), (4), (5) and (9) and methods used for their proof to obtain statistically important estimates for the tail probabilities of the $t-$statistic and the Hotelling T^2 statistic. Dufour and Hallin (1993) performed numerical comparisons of bounds of the type (3)-(4), (6) and (7) and showed that estimates (6) and (7) are substantially superior to their competitors. The authors also discussed applications of the Eaton-type bounds to one-sample permutation $t-$tests, permutation $t-$tests against regression and against first-order autocorrelation and to testing procedures based on linear signed rank statistics.

It is of interest to note here a relation of probability inequalities (1) and (2)-(7) to the finding by Loretan and Phillips (1994, Table 1) that, for typical test sizes, the critical values of the sample split prediction test for covariance stationarity of heavy-tailed time series are lower than in the standard case of time series with innovations having forth moment. For example, from inequality (4) it follows that if $X_1, ..., X_n$ are independent symmetric r.v.'s (not all degenerate), then

$$P\left(\sum_{i=1}^{n} X_i / \left(\sum_{i=1}^{n} X_i^2\right)^{1/2} > x\right) \leq 1 - \Phi(x - 1.5/x),$$

$x > 0$, that implies, in particular, that

$$P\left(\left(\int_0^1 dU^s_{\gamma/2}\right)^{-1} U^s_{\gamma/2}(1) > x\right) \leq 1 - \Phi(x - 1.5/x),$$

where $U^s_{\gamma/2}(1)$ is a symmetric stable process with characteristic exponent $\gamma/2$, $0 < \gamma < 4$. The latter inequality implies, e.g., that the critical values z_α of the sample split tests of size $\alpha\%$ for time series with innovations having Pareto-type tail behavior with tail index γ, $0 < \gamma < 4$, which involve convergence to $\left(\int_0^1 dU^s_{\gamma/2}\right)^{-1} U^s_{\gamma/2}(1)$ are dominated by the quantities $(q_\alpha + \sqrt{q_\alpha^2 + 6})/2$, where q_α is the $(1-\alpha)\%-$quantile of the standard normal distribution: $\Phi(q_\alpha) = 1 - \alpha$. The results in the present paper provide, in particular, extensions of the latter results to the case of multiple stochastic integrals of stable processes.

2. Main inequalities

2a. Inequalities for moving averages and their self-normalized and Studentized versions in independent r.v.'s

Let, as before, Z be the standard normal r.v., $\phi(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}$, and let $\Phi(u) = \int_{-\infty}^{u} \phi(t)dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-\frac{t^2}{2}} dt$. Let $c_i \in \mathbf{R}$, $r_{ki} \in \{0, 1\}$, $k = 1, ..., i-1$, $i = 1, ..., n$, and let $X_1, \ldots, X_n$ be a sample of independent r.v.'s. Consider the random polynomials

$$V_n = \sum_{i=1}^{n} c_i X_1^{r_{1i}} ... X_{i-1}^{r_{i-1,i}} X_i \tag{10}$$

and their self-normalized versions

$$W_n = \sum_{i=1}^{n} c_i X_1^{r_{1i}} ... X_{i-1}^{r_{i-1,i}} X_i / \left(\sum_{i=1}^{n} c_i^2 X_1^{2r_{1i}} ... X_{i-1}^{2r_{i-1,i}} X_i^2\right)^{1/2}. \tag{11}$$

It is easy to see that the class of the above polynomials V_n includes the generalized moving averages $V_n^{mov} = \sum_{i=1}^{n} c_i X_{i+h_1} X_{i+h_2} \ldots X_{i+h_m}$, $0 \leq h_1 < ... < h_m$ (in r.v.'s $X_1, ..., X_{n+h_m}$) arising in a number of problems in econometrics, in particular, the sample auto-covariances $(1/n)\sum_{i=1}^{n} X_i X_{i-1}$ and the sample cross-moments $(1/n)\sum_{i=1}^{n} X_{i+h_1} X_{i+h_2} \ldots X_{i+h_m}$, $0 \leq h_1 < \ldots < h_m$.

The following theorems give sharp generalizations of Hoeffding- and Eaton-Edelman-Pinelis-Dufour-Hallin estimates (1) and (2)-(7) to the case of the statistics V_n and W_n. The inequalities for the self-normalized statistics W_n hold under the only assumption of symmetry of the innovations. E.g., no assumptions on boundedness of the r.v.'s or finiteness of their moments are needed. This property is central in the case of observations coming from a heavy-tailed population. The results imply, essentially, that when applying the probability estimates for sums of dependent r.v.'s formulated at the end of the previous section, in the case of the random polynomials V_n and their self-normalized versions W_n, one can drop the terms accounting for dependence among the summands $c_i X_1^{r_{1i}} ... X_{i-1}^{r_{i-1,i}} X_i$.

Theorem 1 *Let $X_1, \ldots, X_n$ be independent r.v.'s such that $EX_i = 0$, $|X_i| \leq d_i \in \mathbf{R}$ (a.s.), $i = 1, \ldots, n$, and let $D^2 = \sum_{i=1}^n c_i^2 d_1^{2r_{1i}} ... d_{i-1}^{2r_{i-1,i}} d_i^2$, then the following inequalities hold for the random polynomials V_n defined in (10):*

$$P(V_n > x) \leq exp(-\frac{x^2}{2D^2}), \tag{12}$$

$x > 0$,

$$P(V_n > x) \leq \frac{1}{2} \frac{Ef(|Z|)}{f(\frac{x}{D})}, \tag{13}$$

$f \in \overline{K}$, $x > 0$,

$$P(V_n > x) \leq \frac{e^3}{9} \frac{\phi(\frac{x}{D})D}{x}, \tag{14}$$

$x > \sqrt{2}D$,

$$P(V_n > x) \leq \frac{2e^3}{9}(1 - \Phi(\frac{x}{D})), \tag{15}$$

$$P(V_n > x) \leq min(1/2, D^2/(2x^2), B(x/D, n)) \leq$$
$$min(1/2, D^2/(2x^2), \inf_{0<u<x/D} \int_u^\infty ((t-u)^3/(\frac{x}{D} - u)^3)\phi(t)dt) \leq 1 - \Phi(\frac{x}{D} - \frac{1.5D}{x}) \tag{16}$$

$x > 0$, where $B(y, n)$ is defined in (8). The same inequalities, with the right-hand side expressions multiplied by 2, hold for $|V_n|$.

Theorem 2 *Let $X_1, \ldots, X_n$ be independent symmetric r.v.'s (not all degenerate), then the following inequalities hold for the self-normalized random polynomials W_n defined in (11):*

$$P(W_n > x) \leq exp(-\frac{x^2}{2}), \tag{17}$$

$x > 0$,

$$P(W_n > x) \leq \frac{1}{2} \frac{Ef(|Z|)}{f(x)}, \tag{18}$$

$f \in \overline{K}$, $x > 0$,

$$P(W_n > x) \leq \frac{e^3}{9} \frac{\phi(x)}{x}, \tag{19}$$

$x > \sqrt{2}D$,

$$P(W_n > x) \le \frac{2e^3}{9}(1 - \Phi(x)), \tag{20}$$

$$P(W_n > x) \le min(1/2, 1/(2x^2), B(x,n)) \le$$
$$min(1/2, 1/(2x^2), \inf_{0<u<x} \int_u^\infty ((t-u)^3/(x-u)^3)\phi(t)dt) \le 1 - \Phi(x - \frac{1.5}{x}) \tag{21}$$

$x > 0$, *where* $B(y,n)$ *is defined in (8). The same inequalities, with the right-hand side expressions multiplied by 2, hold for* $|W_n|$.

The following theorem gives analogues of Hoeffding and Pinelis-Dufour-Hallin estimates (1) and (7) for the tail probabilities of $t-$statistics in the random polynomials V_n (Studentized random polynomials) that can be applied in testing serial independence of observations (see Section 3). Similar analogues of other exponential inequalities for sums of independent r.v.'s hold as well. The results refine and generalize those obtained by Edelman (1986, 1990) and Pinelis (1994). As in the case of the self-normalized random polynomials W_n, the estimates for the Studentized polynomials hold under the minimal assumption of symmetry of the underlying r.v.'s.

Theorem 3 *Let* $X_1, \dots, X_n$ *be independent symmetric r.v.'s (not all degenerate),* V_n *be as in (10) and let* $\overline{V}_n = (1/n)V_n$, $s^2 = \sum_{i=1}^n (c_i X_{i+h_1} X_{i+h_2} \dots X_{i+h_m} - \overline{V}_n)^2/(n-1)$, *then*

$$P(\sqrt{n}\,\overline{V}_n/s_n > x) \le exp(-nx^2/[2(n-1+x^2)]), \tag{22}$$

$$P(\sqrt{n}\,\overline{V}_n/s_n > x) \le min(1/2, \frac{1+(x^2-1)/n}{2x^2}, B(\frac{x}{(1+(x^2-1)/n)^{1/2}}, n)), \tag{23}$$

$x > 0$, *where* $B(y,n)$ *is defined in (8). The same inequalities, with the right-hand side expressions multiplied by 2, hold for* $|\sqrt{n}\,\overline{V}_n/s_n|$.

2b. Extensions of the results in Sections 1 and 2a to the case of dependent r.v.'s through measures of dependence

Hoeffding (1963), Arcones and Giné (1993) and Eichelsbacher (2000) obtained generalizations of inequality (1) to the case of $U-$statistics. Giné, Latała and Zinn (2000) obtained Bernstein-type inequalities for $U-$statistics.

Several authors (e.g., McConnell and Taqqu (1986), Krakowiak and Szulga (1986), Kwapien and Woyczynski (1992), Szulga (1998) and references therein) have focused on the study of properties of the class of multilinear forms that includes, as a particular case, generalized moving averages, and their applications. In recent years, there has also been increasing interest in the study of sums of multilinear forms, partly because these types of r.v.'s represent a special, but important case of infinite-degree $U-$statistics and are related to the study of long range dependence (cf. Heilig and Nolan (2001)) and moving average processes (e.g., Embrechts, Kluppelberg and Mikosch (1997) and Ho and Hsing (1997)). For example, sums of multilinear forms

$$\sum_{r=0}^{p} d_r \sum_{n=1}^{N} \sum_{1\le j_1<...<j_r<\infty} \prod_{l=1}^{r} c_{j_l} X_{n-j_l},$$

where $c_i \in \mathbf{R}$, $i = 1, 2, \ldots,$ X_i are independent r.v.'s, the function K belongs to a general class of measurable functions, d_r are coefficients depending on K and the distributions of X_i's, represent a stochastic Taylor expansion for functionals of moving averages $\sum_{k=1}^{N} K(\sum_{i=1}^{\infty} c_i X_{n-i})$ important in the study of long-range dependence (see Ho and Hsing (1997)). de la Peña, Ibragimov and Sharakhmetov (2003) obtained sharp moment inequalities for sums of multilinear forms. de la Peña and Zamfirescu (2002) proved domination inequalities for moving averages.

Recently, de la Peña, Ibragimov and Sharakhmetov (2002) obtained sharp estimates for tail probabilities and expected values of statistics in dependent r.v.'s in terms of measures of dependence of the r.v.'s.

Let $\phi^2_{Y_1,...,Y_n}$ and $\delta_{Y_1,...,Y_n}$ denote the following measures of dependence for absolutely continuous or discrete r.v.'s $Y_1, ..., Y_n$ with the one-dimensional distribution functions $F_k(y_k)$, $k = 1, ..., n$, and the joint distribution function $F(y_1, ..., y_n)$:

$$\phi^2_{Y_1,...,Y_n} = \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} \frac{(dF(y_1, ..., y_n))^2}{dF_1(y_1)...dF_n(y_n)} - 1 = \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} (\frac{dF(y_1, ..., y_n)}{dF_1(y_1)...dF_n(y_n)})^2 dF_1(y_1)...dF_n(y_n) - 1$$

(multivariate analog of Pearson's ϕ^2 coefficient),

$$\delta_{Y_1,...,Y_n} = \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} log(\frac{dF(Y_1, ..., Y_n)}{dF_1(y_1)...dF_n(y_n)}) dF(y_1, ..., y_n)$$

(relative entropy), where the integrals are in the sense of Lebesque-Stiltjes and $\frac{dF(y_1,...,y_n)}{dF_1(y_1)...dF_n(y_n)}$ is to be taken to be 0 if $dF_1(y_1)...dF_n(y_n) = 0$ in the former case and to be 1 if $dF_1(y_1)...dF_n(y_n) = 0$ in the latter case. In the case of absolutely continuous r.v.'s $Y_1, ..., Y_n$ the multivariate measures $\delta_{Y_1,...,Y_n}$ and $\phi^2_{Y_1,...,Y_n}$ were introduced by Joe (1987, 1989). In the bivariate case, the measures $\phi^2_{Y_1,Y_2}$ and δ_{Y_1,Y_2} are commonly known as Pearson's ϕ^2 coefficient and the mutual information between Y_1 and Y_2, respectively. If $(Y_1, ..., Y_n)' \sim N(\mu, \Sigma)$, then (see Joe (1989)) $\phi^2_{Y_1,...,Y_n} = |R(2I_n - R)|^{-1/2} - 1$, where I_n is the $n \times n$ identity matrix, provided that the correlation matrix R corresponding to Σ has maximum eigenvalue less than 2 and is infinite otherwise ($|A|$ denotes the determinant of a matrix A). In addition to that, if in the above case $diag(\Sigma) = (\sigma_1^2, ..., \sigma_n^2)$, then $\delta_{Y_1,...,Y_n} = -.5log(|\Sigma|/\prod_{i=1}^{n} \sigma_i^2)$. In the case of two normal r.v.'s Y_1 and Y_2 with correlation coefficient ρ, $(\phi^2_{Y_1,Y_2}/(1 + \phi^2_{Y_1,Y_2}))^{1/2} = (1 - exp(-2\delta_{Y_1,Y_2}))^{1/2} = |\rho|$.

de la Peña, Ibragimov and Sharakhmetov (2002) showed that the following complete decoupling estimates hold for the tail probabilities of arbitrary statistics $h(Y_1, ..., Y_n)$ in r.v.'s $Y_1, ..., Y_n$:

$$P(h(Y_1, ..., Y_n) > x) \le P(h(\xi_1, ..., \xi_n) > x) + \phi_{Y_1,...,Y_n} (P(h(\xi_1, ..., \xi_n) > x))^{1/2}, \tag{24}$$

$$P(h(Y_1, ..., Y_n) > x) \le \left(1 + \phi^2_{Y_1,...,Y_n}\right)^{1/2} (P(h(\xi_1, ..., \xi_n) > x))^{1/2}, \tag{25}$$

$$P(h(Y_1, ..., Y_n) > x) \le (e-1)P(h(\xi_1, ..., \xi_n) > x) + \delta_{Y_1,...,Y_n}, \tag{26}$$

where $\xi_1, ..., \xi_n$ denote independent copies of the dependent r.v.'s $Y_1, ..., Y_n$. The latter results and inequalities (1) and (2)-(7) for sums of independent r.v.'s imply corresponding sharp probability inequalities for sums of dependent r.v.'s. The following exact analogues of Hoeffding's inequality (1) and Pinelis-Dufour-Hallin estimate (7) for dependent r.v.'s hold: If $Y_1, ..., Y_n$ are r.v.'s with $EY_i = 0$, $i = 1, \ldots, n$, such that $|Y_i| \le d_i \in \mathbf{R}$ (a.s.), $i = 1, \ldots, n$, then

$$P\left(\sum_{i=1}^{n} Y_i > x\right) \le exp\left(-\frac{x^2}{2D^2}\right) + \phi_{Y_1,...,Y_n} exp\left(-\frac{x^2}{4D^2}\right), \tag{27}$$

$$P\left(\sum_{i=1}^{n} Y_i > x\right) \le (1 + \phi^2_{Y_1,...,Y_n})^{1/2} exp\left(-\frac{x^2}{4D^2}\right), \tag{28}$$

$$P\left(\sum_{i=1}^{n} Y_i > x\right) \le (e-1) exp\left(-\frac{x^2}{2D^2}\right) + \delta_{Y_1,...,Y_n}, \tag{29}$$

$$P\left(\sum_{i=1}^{n} Y_i > x\right) \le min\left(1/2, D^2/(2x^2), B(x/D, n)\right) + \phi_{Y_1,...,Y_n} min\left(1/\sqrt{2}, D/(\sqrt{2}x), (B(x/D, n))^{1/2}\right), \tag{30}$$

$$P\left(\sum_{i=1}^{n} Y_i > x\right) \le \left(1 + \phi^2_{Y_1,...,Y_n}\right) min\left(1/\sqrt{2}, D/(\sqrt{2}x), (B(x/D, n))^{1/2}\right), \tag{31}$$

$$P\left(\sum_{i=1}^{n} Y_i > x\right) \le (e-1) min\left(1/2, D^2/(2x^2), B(x/D, n)\right) + \delta_{Y_1,...,Y_n}, \tag{32}$$

$x > 0$, where $B(y, n)$ is defined in (8).

From inequalities (24)-(26) it follows that estimates similar to those in Theorems 1- 3 and involving the measures of dependence ϕ^2 and δ hold also for the random polynomials V_n and their self-normalized and Studentized analogues W_n and $\sqrt{n}\,\overline{V}_n/s_n$ in dependent r.v.'s. For example, the following theorems give analogues of inequality (7) for V_n, W_n and $\sqrt{n}\,\overline{V}_n/s_n$.

In the inequalities throughout the rest of the paper, the extremal cases of the estimates $+\infty \le +\infty$, $-\infty \le +\infty$ and $-\infty \le -\infty$ are considered to be valid inequalities; we, therefore, usually do not include assumptions on finiteness of moments of the summand r.v.'s that ensure finiteness of moments of sums of the r.v.'s into formulations of the results.

Theorem 4 *Let $X_1,\ldots,X_n$ be absolutely continuous or discrete mean-zero dependent r.v.'s such that $|X_i| \le d_i \in \mathbf{R}$, $i = 1,...,n$, then the following inequalities hold for the random polynomials V_n defined in (10) in r.v.'s $X_1,\ldots,X_n$:*

$$P(V_n > x) \le min(1/2, D^2/(2x^2), B(x/D,n)) + \phi_{X_1,\ldots,X_n} min(1/\sqrt{2}, D/(\sqrt{2}x), (B(x/D,n))^{1/2}), \tag{33}$$

$$P(V_n > x) \le (1+\phi^2_{X_1,\ldots,X_n})^{1/2} min(1/\sqrt{2}, D/(\sqrt{2}x), (B(x/D,n))^{1/2}),$$

$$P(V_n > x) \le (e-1) min(1/2, D^2/(2x^2), B(x/D,n)) + \delta_{X_1,\ldots,X_n},$$

$x > 0$, *where* $D^2 = \sum_{i=1}^n c_i^2 d_1^{2r_{1i}} \ldots d_{i-1}^{2r_{i-1,i}} d_i^2$.

Theorem 5 *Let $X_1,\ldots,X_n$ be absolutely continuous or discrete symmetric dependent r.v.'s (not all degenerate), then the following inequalities hold for the self-normalized random polynomials W_n defined in (11) in r.v.'s $X_1,\ldots,X_n$:*

$$P(W_n > x) \le min(1/2, 1/(2x^2), B(x,n)) + \phi_{X_1,\ldots,X_n} min(1/\sqrt{2}, 1/(\sqrt{2}x), (B(x,n))^{1/2}), \tag{34}$$

$$P(W_n > x) \le (1+\phi^2_{X_1,\ldots,X_n})^{1/2} min(1/\sqrt{2}, 1/(\sqrt{2}x), (B(x,n))^{1/2}),$$

$$P(W_n > x) \le (e-1) min(1/2, 1/(2x^2), B(x,n)) + \delta_{X_1,\ldots,X_n},$$

$x > 0$.

Theorem 6 *Let $X_1,\ldots,X_n$ be symmetric r.v.'s (not all degenerate) and let*

$$\overline{V}_n = (1/n)V_n, \;\; s^2 = \sum_{i=1}^n (c_i X_{i+h_1} X_{i+h_2} \ldots X_{i+h_m} - \overline{V}_n)^2/(n-1).$$

It follows that

$$P(\sqrt{n}\,\overline{V}_n/s_n > x) \le min\left\{1/2, (n-1+x^2)/(2nx^2), B(n^{1/2}x/(n-1+x^2)^{1/2}, n)\right\} + \phi_{Y_1,\ldots,Y_n} min\left\{1/\sqrt{2}, (n-1+x^2)^{1/2}/((2n)^{1/2}x), \left(B(n^{1/2}x/(n-1+x^2)^{1/2}, n)\right)^{1/2}\right\}, \tag{35}$$

$$P(\sqrt{n}\,\overline{V}_n/s_n > x) \le (1+\phi^2_{Y_1,\ldots,Y_n})^{1/2} min\left\{1/\sqrt{2}, (n-1+x^2)^{1/2}/((2n)^{1/2}x), \left(B(n^{1/2}x/(n-1+x^2)^{1/2}, n)\right)^{1/2}\right\},$$

$$P(\sqrt{n}\,\overline{V}_n/s_n > x) \le (e-1) min\left\{1/2, (n-1+x^2)/(2nx^2), B(n^{1/2}x/(n-1+x^2)^{1/2}, n)\right\} + \delta_{Y_1,\ldots,Y_n},$$

$x > 0$. *The same inequalities, with the right-hand side expressions multiplied by 2, hold for* $|\sqrt{n}\,\overline{V}_n/s_n|$.

Note again that essentially only the condition on symmetry of the r.v.'s $X_1, ..., X_n$ is required for the estimates in Theorems 5 and 6 to hold. It is also emphasized here that bounds (33)-(35) for the statistics V_n, W_n and $\sqrt{n}\, V_n/s_n$ in dependent r.v.'s become exactly the analogues of the Pinelis-Duffour-Hallin inequalities given by (16), (21) and (23) in the case $\phi^2_{X_1,...,X_n} = 0$, that is, in the case of the statistics in independent r.v.'s.

2c. Sharp moment inequalities for moving averages and their applications

The following result extends inequality (9) to the case of the random polynomials V_n in independent mean-zero r.v.'s. The results give sharp generalizations of the Khintchine-Marcinkiewicz-Zygmund-type inequalities obtained, in the case of sums of independent r.v.'s, by Eaton (1970, 1974), Pinelis (1994) and Figiel, Hitczenko, Johnson, Schechtman and Zinn (1997).

Again, let K be the class of twice differentiable even functions $f : \mathbf{R} \to \mathbf{R}$ such that f'' is nonnegative and convex. In what follows, ϵ, ϵ_t, $t \in \{..., -2, -1, 0, 1, 2, ...\}$, denote independent symmetric Bernoulli r.v.'s.

Theorem 7 *Let $f \in K$, $X_1, \ldots, X_n$ be independent r.v.'s such that $EX_i = 0$, $|X_i| \le 1$, $i = 1, \ldots, n$, and let $\tilde{X}_1, \ldots, \tilde{X}_n$ be independent symmetric r.v.'s such that $E\tilde{X}_i^2 = 1$, $i = 1, \ldots, n$, then*

$$Ef(V_n) \le Ef(\sum_{i=1}^n c_i \epsilon_i) \le Ef(\sum_{i=1}^n c_i \tilde{X}_i). \tag{36}$$

If, in addition to the above, $\sum_{i=1}^n c_i^2 = 1$, then

$$Ef(V_n) \le Ef(Z). \tag{37}$$

As de la Peña, Ibragimov and Sharakhmetov (2003) showed, the best constants in the Khintchine-Marcinkie- wicz-Zygmund inequalities for powers of generalized moving averages in symmetric r.v.'s are the same as in the case of sums of independent r.v.'s. According to the following theorem, the same result holds for the Khintchine-Marcinkiewicz-Zygmund inequalities as well as for Dharmadhikari-Jogdeo-type (see Dharmadhikari and Jogdeo (1969)) inequalities for the random polynomials V_n.

Theorem 8 *The best constants $A_1^*(t, m)$, $B_1^*(t, m)$, $A_2^*(t)$ and $B_2^*(t)$ in the following Khintchine-Marcinkiewicz-Zygmund inequalities*

$$A(t,m)E(\sum_{i=1}^n c_i^2 X_1^{2r_{1i}} ... X_{i-1}^{2r_{i-1,i}} X_i^2)^{t/2} \le E|\sum_{i=1}^n c_i X_1^{r_{1i}} ... X_{i-1}^{r_{i-1,i}} X_i|^t \le$$
$$B(t,m)E(\sum_{i=1}^n c_i^2 X_1^{2r_{1i}} ... X_{i-1}^{2r_{i-1,i}} X_i^2)^{t/2} \tag{38}$$

for all independent symmetric r.v.'s $X_1, \ldots, X_n$ with finite $t-$th moment, $t > 0$, are given by $A^*(t, m) = 2^{t/2-1}$, $0 < t \le t_0$, $A^*(t, m) = E|Z|^t$, $t_0 \le t \le 2$, $A^*(t, m) = 1$, $t \ge 2$, $B^*(t, m) = 1$, $0 < t \le 2$, $B^*(t, m) = E|Z|^t$,

$t \geq 2$, *where* t_0 *is the nontrivial solution of the equation* $\Gamma((t_0+1)/2) = \Gamma(3/2)$, $\Gamma(x)$ *is the Gamma-function:* $\Gamma(x) = \int_0^{+\infty} t^{x-1}e^{-t}dt$.

The best constant $C^*(t,m)$ *in the following Dharmadhikari-Jogdeo-type inequality*

$$E|\sum_{i=1}^{n} c_i X_1^{r_{1i}}...X_{i-1}^{r_{i-1,i}} X_i|^t \leq C(t,m) n^{t/2-1} \sum_{i=1}^{n} |c_i|^t E|X_1|^{tr_{1i}}...E|X_{i-1}|^{tr_{i-1,i}} E|X_i|^t \tag{39}$$

for all independent symmetric r.v.'s $X_1, \ldots, X_n$ *with finite* $t-th$ *moment,* $t \geq 2$, *is given by* $C^*(t,m) = E|Z|^t$.

Using estimate (39) and Hölder's inequality, we obtain the following result that generalizes the results obtained in Anderson (1993) and Anderson and Chen (1996) and gives a sharp estimate for the greatest order, in n, that moments of generalized moving averages $\sum_{i=1}^n X_{i+h_1}...X_{i+h_m}$ and sample cross-moments $(1/n)\sum_{i=1}^n X_{i+h_1}...X_{i+h_m}$ in independent mean-zero r.v.'s X_i can attain. As usual, the notation $a_n = O(b_n)$ for two nonnegative sequences (a_n) and (b_n), $n \geq 1$, means that $a_n \leq Cb_n$, $n \geq 1$, for some constant C that does not depend on n.

If $t_1, ..., t_k > 2$, $t = \sum_{s=1}^k t_s$, $0 \leq h_1^{(s)} < ... < h_m^{(s)}$, $s = 1, ..., k$; $X_1, \ldots, X_{n+h_m}$, where $h_m = \max_{s=1,k} h_m^{(s)}$, are independent identically distributed r.v.'s with $EX_1 = 0$ and $E|X_1|^t < \infty$, then

$$E \prod_{s=1}^{k} |\sum_{i=1}^{n} X_{i+h_1^{(s)}} X_{i+h_2^{(s)}} \ldots X_{i+h_m^{(s)}}|^{t_s} = O(n^{t/2}).$$

The following theorem gives an estimate for the rate of convergence in the central limit theorem for moments of random polynomials V_n (introduced in Section 2a) that generalizes the classical results of von Bahr (1965) and Michel (1976) for sums of independent r.v.'s.

Theorem 9 *If* $3 \leq t < 4$, $X_1, ..., X_n$ *are independent identically distributed symmetric r.v.'s with* $EX_1^2 = 1$, $E|X_1|^t < \infty$, *then*

$$|E|n^{-1/2}V_n|^t - E|Z|^t| = O(n^{1-t/2}). \tag{40}$$

3. Applications of the results for moving averages in hypothesis testing

In the present and the next sections, we deal with applications of the estimates considered above to several problems in statistical inference. The applications are motivated by the fact that, as discussed before, the class of random polynomials V_n includes the generalized moving averages and sample cross-moments frequently arising in statistical and econometric studies.

3a. Permutation tests against serial correlation and tests for independence

Consider the problems of testing that cross-moments of a mean-zero stationary time series X_t, $t = 0, 1, 2, ...$, equal zero: $EX_{h_1}X_{h_2}...X_{h_m} = 0$, $0 \leq h_1 < ... < h_m$, e.g., that the r.v.'s are uncorrelated $EX_1X_2 = 0$ (the setup appears in a natural way in testing joint independence of the r.v.'s). These problems, arise, in particular, in the problem of testing $H_0 : \rho = 0$ against $H_A : \rho > 0$ in the first-order autoregressive model

$$X_t = \rho X_{t-1} + u_t, \tag{41}$$

$t = 0, 1, ..., n$, where $u_0, u_1, ..., u_n$ are independent random disturbances with possibly nonidentical distributions symmetric about 0 (one evidently has $EX_{h_1}X_{h_2}...X_{h_m} = 0$, $0 \le h_1 < ... < h_m$, e.g., $EX_1X_2 = 0$ under H_0). Testing H_0 can be essentially reduced to testing that the mean of the series $Y_t = X_{t+h_1}X_{t+h_2}...X_{t+h_m}$, $t = 1, 2, ...$, (respectively, $Y_t = X_tX_{t+1}$, $t = 1, 2, ...$) is zero. As in the above standard setup, the testing procedures for these problems can be based, therefore, on the $t-$statistics $\sqrt{n}\,\overline{V}^{(1)}_{n,m}/s^{(1)}_{n,m}$ (the superscript (1) refers to the moving average form of the statistics), where

$$\overline{V}^{(1)}_{n,m} = (1/n)\sum_{i=1}^{n} X_{i+h_1}X_{i+h_2}...X_{i+h_m}$$

and

$$(s^{(1)}_{n,m})^2 = \sum_{i=1}^{n}\left(X_{i+h_1}X_{i+h_2}...X_{i+h_m} - \overline{V}^{(1)}_{n,m}\right)^2/(n-1).$$

Evidently, under the null hypothesis, the tail probabilities of the Studentized generalized moving averages $\overline{V}^{(1)}_{n,m}/s^{(1)}_{n,m}$, e.g., the Studentized sample auto-covariances

$$\overline{V}^{(1)}_{n,1}/s^{(1)}_{n,1} = \sqrt{(n-1)/n}\sum_{i=1}^{n} X_iX_{i+1}/\left(\sum_{i=1}^{n}\left(X_iX_{i+1} - 1/n\sum_{i=1}^{n} X_iX_{i+1}\right)^2\right)^{1/2},$$

satisfy the inequalities in Theorem 3. This implies that, when applying the above testing procedures in the latter setup one can in fact drop the terms accounting for dependence among the summands Y_t in estimates (48)-(53) (and similar generalizations of other estimates for the tail probabilities of the $t-$statistics in independent r.v.'s). We then have that

$$P(\sqrt{n}\,\overline{V}^{(1)}_{n,m}/s^{(1)}_{n,m} > x) \le 1 - \Phi\left(n^{1/2}x/(n-1+x^2)^{1/2} - 1.5(n-1+x^2)^{1/2}/(n^{1/2}x)\right), \tag{42}$$

$$P(\sqrt{n}\,\overline{V}^{(1)}_{n,m}/s^{(1)}_{n,m} > x) \le \frac{2e^3}{9}\left(1 - \Phi\left(n^{1/2}x/(n-1+x^2)^{1/2}\right)\right), \tag{43}$$

$$P(\sqrt{n}\,\overline{V}^{(1)}_{n,m}/s^{(1)}_{n,m} > x) \le min\left\{1/2, (n-1+x^2)/(2nx^2), B(n^{1/2}x/(n-1+x^2)^{1/2}, n)\right\}, \tag{44}$$

where $B(y, n)$ is defined in (8). Consequently, one can use, in particular, the following conservative critical region for the test $H_0 : \rho = 0$ against $H_A : \rho > 0$ with level α : $\sqrt{n}\,\overline{V}^{(1)}_{n,m}/s_{n,m} > y^{(i)}_\alpha$, $i = 1, 2, 3$, where $y^{(i)}_\alpha$, $i = 1, 2, 3$, are such that

$$1 - \Phi\left(n^{1/2}y^{(1)}_\alpha/(n-1+(y^{(1)}_\alpha)^2)^{1/2} - 1.5(n-1+(y^{(1)}_\alpha)^2)^{1/2}/(n^{1/2}y^{(1)}_\alpha)\right) < \alpha, \tag{45}$$

$$\frac{2e^3}{9}\left(1-\Phi\left(n^{1/2}y_\alpha^{(2)}/(n-1+(y_\alpha^{(2)})^2)^{1/2}\right)\right)<\alpha, \tag{46}$$

$$min\left\{1/2,(n-1+(y_\alpha^{(3)})^2)/(2n(y_\alpha^{(3)})^2),B(n^{1/2}y_\alpha^{(3)}/(n-1+(y_\alpha^{(3)})^2)^{1/2},n)\right\}<\alpha, \tag{47}$$

and $B(y,n)$ is defined in (8).

One can also consider the sign versions of the above tests based on the fact that, under H_0, $Esign(X_{h_1})sign(X_{h_2})...sign(X_{h_m})=0$, $0\le h_1<...<h_m$ (in particular, $Esign(X_1)sign(X_2)=0$), where $sign(X_t)$ is the sign of X_t defined by $sign(X_t)=1$ if $X_t>0$, $sign(0)=0$ and $sign(X_t)=-1$ otherwise. The latter testing procedures can be based on the statistics $\sqrt{n}\,\overline{V}_{n,m}^{(1,sign)}/s_{n,m}^{(1,sign)}$, where

$$\overline{V}_{n,m}^{(1,sign)}=(1/n)\sum_{i=1}^n sign(X_{i+h_1})sign(X_{i+h_2})...sign(X_{i+h_m})$$

and

$$(s_{n,m}^{(1,sign)})^2=\sum_{i=1}^n\left(sign(X_{i+h_1})sign(X_{i+h_2})...sign(X_{i+h_m})-\overline{V}_{n,m}^{(1,sign)}\right)^2/(n-1).$$

Evidently, under the null hypothesis, the sign versions $\sqrt{n}\overline{V}_{n,m}^{(1,sign)}/s_{n,m}^{(1,sign)}$ of the statistics $\overline{V}_{n,m}^{(1)}/s_{n,m}^{(1)}$ satisfy the same inequalities as above.

Note that the conservative tests based on the above Studentized statistics $\sqrt{n}\,\overline{V}_{n,m}^{(1)}/s_{n,m}^{(1)}$ and their sign versions $\sqrt{n}\,\overline{V}_{n,m}^{(1,sign)}/s_{n,m}^{(1,sign)}$ are equivalent to the tests based on the self-normalized moving averages

$$W_n^{(1)}=\sum_{i=1}^n X_{i+h_1}X_{i+h_2}...X_{i+h_m}/\left(\sum_{i=1}^n X_{i+h_1}^2X_{i+h_2}^2...X_{i+h_m}^2\right)^{1/2}$$

and their sign versions

$$W_n^{(1,sign)}=\sum_{i=1}^n sign(X_{i+h_1})sign(X_{i+h_2})...sign(X_{i+h_m})/$$

$$\left(\sum_{i=1}^n(sign(X_{i+h_1})sign(X_{i+h_2})...sign(X_{i+h_m}))^2\right)^{1/2}.$$

In the case of the r.v.'s $X_1,...,X_{n+h_m}$ such that $P(X_k=0)=0$, $k=1,...,n+h_m$, the sign versions of the tests are evidently equivalent to those based on the statistics

$$n^{-1/2}\sum_{i=1}^n sign(X_{i+h_1})sign(X_{i+h_2})...sign(X_{i+h_m}).$$

The statistics $W_n^{(1)}$ and $W_n^{(1,sign)}$ satisfy the inequalities in Theorem 2 which imply conservative critical regions for the statistics analogous to those above for the Studentized statistics $V_n^{(1)}$ and $V_n^{(1,sign)}$.

One should note that, in the case $m=1$, the tests based on the Studentized sample auto-covariance $\sqrt{n}\overline{V}_{n,1}^{(1)}/s_{n,1}^{(1)}$ and estimate (44) (that is , the tests with the critical regions determined by (47)) are essentially equivalent to the permutation tests against first-order auto-regression based on the non-uniform estimates for the first-order auto-correlation coefficient proposed by Dufour and Hallin (1993). One should also emphasize here that applications of the inequalities of type (44) (e.g., the tests based on the critical regions (47)) involve the problems of determining numerically the minima of the functions in the definition of $B(y, n)$ (see Dufour and Hallin (1993)) that are usually computationally intensive. On the other hand, the applications of the uniform bounds such as (42) and (43) are less computationally intensive although they yield more conservative tests than those based on (44).

If in the model (41) the disturbances u_t are dependent, one can use conservative critical regions for the tests implied by Theorem 3 and estimates (24)-(26) for statistics in dependent r.v.'s. In the latter case the statistics $\sqrt{n}\overline{V}_{n,m}^{(1)}/s_{n,m}^{(1)}$ satisfy inequalities analogous to (48)-(53).

The conservative tests based on the above statistics $\sqrt{n}\,\overline{V}_{n,m}^{(1)}/s_{n,m}^{(1)}$ and their sign versions $\sqrt{n}\,\overline{V}_{n,m}^{(1,sign)}/s_{n,m}^{(1,sign)}$ can also be applied in the problems of testing for joint independence in a sample of r.v's $X_1, ..., X_n$.

3b. Monte-Carlo study

To illustrate the statistical properties of the testing procedures described above, we present in this section the results of a Monte-Carlo study of size and power of two of the conservative tests against serial dependence in model (41). We focus on the conservative two-sided tests of the hypothesis $H_0 : \rho = 0$ based on the uniform bounds (42) and (43) for Studentized sample covariances (Studentized moving averages of the first order). The test statistics, similar to the setup considered in the previous section, are $\sqrt{n}\,\overline{V}_n^{(mov)}/s_n^{(mov)}$ and their sign versions $\sqrt{n}\,\overline{V}_n^{(mov,sign)}/s_n^{(mov,sign)}$, where

$$\overline{V}_n^{(mov)} = (1/n)\sum_{i=1}^{n} X_i X_{i-1},$$

$$(s_n^{(mov)})^2 = \sum_{i=1}^{n}\left(X_i X_{i-1} - \overline{V}_n^{(mov)}\right)^2/(n-1),$$

and

$$\overline{V}_n^{(mov,sign)} = (1/n)\sum_{i=1}^{n} sign(X_i)sign(X_{i-1}),$$

$$(s_n^{(mov)})^2 = \sum_{i=1}^{n}\left(sign(X_i)sign(X_{i-1}) - \overline{V}_n^{(mov,sign)}\right)^2/(n-1).$$

Clearly, the test statistic satisfies the above bounds (42) and (43) under H_0 and their two-sided versions, with the right-hand sides multiplied by 2.

We consider the following distributions for the errors u_t in model (41):

D1, $N(0, 1)$, the standard normal distribution;

D2, $0.95N(0,1)+0.05N(0,25) = N(0,1.44)$, a variance mixture of two normal distributions (as in So Shin (2001));

D3, the standard normal distribution for one-half of the sample and a normal distribution with mean 0, variance 16 for the other half (similarly to Dufour and Hallin (1991))

D4, the standard Cauchy distribution;

D5, the standard normal distribution for one-half of the sample and the standard Cauchy distribution for the other half.

The distributions D2 (see So and Shin (2001)) and distributions D3-D5 have heavier tails than in the case of the normal distributions. The samples in D3 and D5 have nonidentical distributions. All calculations reported below were performed using the GAUSS (2001) statistical package.

Table 1 gives the results on the size properties of the conservative tests of level 0.05 based on $\sqrt{n}\,\overline{V}_n^{(mov)}/s_n^{(mov)}$ and estimates (42) and (43) for the sample sizes $n = 20, 50, 100, 200$ and 500. Table 2 gives the results on the empirical sizes of the sign analogues of the latter tests based on $\sqrt{n}\,\overline{V}_n^{(mov,sign)}/s_n^{(mov,sign)}$. The percentages of rejections were computed by simulation with 10,000 replications. For illustrative purposes, for the cases D1 and D2 of the standard normal distribution and the variance mixture of normals, we also present the results on the size properties of the conventional test against serial dependence in model (41) based on the first-order autocorrelation coefficient $r_n = \sum_{i=1}^n X_i X_{i-1}/\sum_{i=1}^n X_i^2$ and the well-known result that $\sqrt{n} r_n \to N(0,1)$ (in distribution) under H_0 if $Eu_t^2 < \infty$. The percentages of rejections in the latter case are reported in the third column for each distribution (the column "limit, autoregression").

[Table 1 around here]

Tables 3a, b and 4a, b below contain the results of a Monte-Carlo study of the power properties of the above conservative tests of level 0.05 against the alternatives $\rho = 0.5$ and $\rho = -0.5$. The sample sizes and the number of replications in the experiment are the same as above. For an illustration, we also present the percentages of rejections in the case of the test based on the normal convergence of the sample autocorrelation.

[Table 2 around here]

[Table 3 around here]

As Tables 1 and 2 illustrate, the conservative tests based on bounds (42) and (43) for the Studentized autocovariances and their sign versions have remarkably good size properties. This feature of the tests is stable to the sample sizes in the model and to the heaviness of the tails of innovations. In fact, the size properties of the tests based on $\sqrt{n}\,\overline{V}_n^{(mov)}/s_n^{(mov)}$ for the heavy-tail Cauchy distribution outperform considerably the cases of the standard normal distribution and the variance mixture of normals. In the case of the distributions D1 and D2, the conservative tests and their sign analogues provide a considerable improvement over the conventional test based on the normal convergence of the first-order auto-correlation coefficient in model (41) with $\rho = 0$ and innovations with a finite second moment. The power properties of the tests based on $\sqrt{n}\,\overline{V}_n^{(mov)}/s_n^{(mov)}$ and (42) and (43), as illustrated by Tables 3a, b, improve considerably with an increase in the sample size in all the five distributional cases. The percentage of rejections for the standard normal distribution and the variance mixture of normals increases from 14-15% to 99-100% as the sample size increases from 20 to 100 and higher. The percentage of rejections for the Cauchy distribution increases from 3% for the sample size 20 to 18-25% in the case of the sample sizes 100-500. For sample sizes greater than 100, the power of the conservative tests for the normal and the mixed normal case is approximately the same as that for the conventional test based on the normal convergence of the sample auto-regression coefficient. According to Tables 4a and 4b, the empirical powers for the case of heavy tailed distributions D4 and D5 improve considerably in the case of the tests based on the sign versions of the statistics $\sqrt{n}\,\overline{V}_n^{(mov,sign)}/s_n^{(mov,sign)}$. For the sign tests, the percentage of rejections for the standard normal distribution and the variance mixture of normals (the distributions D1 and D2) increases from 8-9% to 98-100% as the sample size increases from 20 to 200 and higher. The empirical powers of the sign versions of the tests for the heavy-tailed distributions even outperform the cases of the distributions D1 and D2, increasing, e.g., in the case of the Cauchy distribution D4, from 14-15% for the sample size 20 to 93-100% in the case of the samples of size 100 and higher.

4. Applications in testing procedures for dependent r.v.'s

4a. Conservative critical regions for nonparametric $t-$tests

Let $Y_1, ..., Y_n$ be r.v.'s with unspecified (possibly nonidentical) one-dimensional distributions symmetric about a common median μ and consider the problem of testing $H_0 : \mu = \mu_0$ against $\mu > \mu_0$. The most widely used test statistic for a problem of this type is the one-sample $t-$statistic

$$T_n = n^{-1/2}\sum_{i=1}^{n}(Y_i - \mu_0)/\left((n-1)^{-1}\sum_{i=1}^{n}(Y_i - \bar{Y})^2\right)^{1/2},$$

where $\bar{Y} = (1/n)\sum_{i=1}^{n} Y_i$. It is usually assumed that the r.v.'s $Y_1, ..., Y_n$ are independent identically distributed normal r.v.'s, in which case T_n follows a $t-$distribution with $n-1$ degrees of freedom under the hypothesis H_0. However the result can no longer be used if the distributions of the Y_i's are unknown, nonidentical or if there is dependence among Y_i's. In this context, bounds for tail probabilities of T_n that hold for all symmetric r.v.'s $Y_1, ..., Y_n$ for different classes of dependent r.v.'s become important.

Using estimates (24)-(26) similarly to Edelman (1986, 1990) and Dufour and Hallin (1991) one can easily derive estimates for $P(T_n > x)$, $x > 0$, in the case of arbitrary absolutely continuous or discrete symmetric r.v.'s $Y_1, ..., Y_n$ with the dependence characteristics $\phi^2_{Y_1,...,Y_n}$ or $\delta_{Y_1,...,Y_n}$. For example, since $\sum_{i=1}^{n}(\sum_{j=1}^{n}(Y_j - \mu_0)^2)^{-1/2})(Y_i - \mu_0) = n^{1/2}T_n/(n-1+T_n^2)^{1/2}$, from inequalities (24)-(26) and estimates (1) and (7) it follows that the following bounds for the tail probabilities of the statistic T_n in r.v.'s $Y_1, ..., Y_n$ hold:

$$P(T_n > x) \leq exp(-nx^2/[2(n-1+x^2)]) + \phi_{Y_1,...,Y_n} exp(-nx^2/[4(n-1+x^2)]), \tag{48}$$

$$P(T_n > x) \leq (1 + \phi^2_{Y_1,...,Y_n})^{1/2} exp(-nx^2/[4(n-1+x^2)]), \tag{49}$$

$$P(T_n > x) \leq exp(-nx^2/[2(n-1+x^2)]) + \delta_{Y_1,...,Y_n}, \tag{50}$$

$$P(T_n > x) \leq min\Big\{1/2, (n-1+x^2)/(2nx^2), B(n^{1/2}x/(n-1+x^2)^{1/2}, n)\Big\} + \phi_{Y_1,...,Y_n} min\Big\{1/\sqrt{2}, (n-1+x^2)^{1/2}/((2n)^{1/2}x), \Big(B(n^{1/2}x/(n-1+x^2)^{1/2}, n)\Big)^{1/2}\Big\}, \tag{51}$$

$$P(T_n > x) \leq (1 + \phi^2_{Y_1,...,Y_n})^{1/2} min\Big\{1/\sqrt{2}, (n-1+x^2)^{1/2}/((2n)^{1/2}x), \Big(B(n^{1/2}x/(n-1+x^2)^{1/2}, n)\Big)^{1/2}\Big\}, \tag{52}$$

$$P(T_n > x) \leq (e-1) min\Big\{1/\sqrt{2}, (n-1+x^2)^{1/2}/((2n)^{1/2}x), \Big(B(n^{1/2}x/(n-1+x^2)^{1/2}, n)\Big)^{1/2}\Big\} + \delta_{Y_1,...,Y_n}, \tag{53}$$

$x > 0$. Let us note that the bounds of the type (48) and (51) become exactly the estimates for $P(T_n > x)$ implied by (1) and (5)-(7) in the independent case, that is, in the case $\phi^2_{Y_1,...,Y_n} = 0$. In particular, estimate (48) becomes exactly the bound for $P(T_n > x)$ in the case of independent r.v.'s $Y_1, ..., Y_n$ obtained by Edelman (1986).

It is interesting to note that the bounds that can be derived using the above approach can be improved in the case of identically distributed, but possibly correlated, normal r.v.'s $X_1, ..., X_n \sim N(\mu, \Sigma)$, $diag(\Sigma) = (\sigma^2, ..., \sigma^2)$, such that the correlation matrix R corresponding to Σ has a maximum eigenvalue less than 2. Namely, using the fact that in the above case, $\phi^2_{X_1,...,X_n} = |R(2I_n - R)|^{-1/2} - 1$ and $\delta_{X_1,...,X_n} = -0.5log(|\Sigma|/\sigma^{2n})$ (see Section 1), where I_n is the $n \times n$ identity matrix, and the statistic T_n follows $t-$ distribution with $n-1$ degrees of freedom in the case of identically distributed independent normal r.v.'s X_i, from (24)-(26) we get that, for $x > 0$,

$$P(T_n > x) \leq P(t_{n-1} > x) + (|R(2I_n - R)|^{-1/2} - 1)^{1/2}[P(t_{n-1} > x)]^{1/2},$$

$$P(T_n > x) \leq |R(2I_n - R)|^{-1/4}[P(t_{n-1} > x)]^{1/2},$$

$$P(T_n > x) \leq (e-1)P(t_{n-1} > x) - 0.5log(|\Sigma|/\sigma^{2n}),$$

where t_{n-1} is a t-distributed r.v. with $n-1$ degrees of freedom. A conservative critical region for the one-sided $t-$test with level α is given by $T_n > y_\alpha$, where y_α is such that

$$min\Big\{P(t_{n-1} > y_\alpha) + (|R(2I_n - R)|^{-1/2} - 1)^{1/2}[P(t_{n-1} > y_\alpha)]^{1/2}, |R(2I_n - R)|^{-1/4}[P(t_{n-1} > y_\alpha)]^{1/2},$$

$$(e-1)P(t_{n-1} > y_\alpha) - 0.5log(|\Sigma|/\sigma^{2n})\big\} < \alpha.$$

4b. Conservative tests of linear hypotheses

An approach similar to that in Section 4a can be applied in other testing procedures. Consider, for example, the linear regression model $y = X\beta + u$, where X is an $n\times k$ full rank scalar matrix, $y \in \mathbf{R}^{n\times 1}$; $\beta \in \mathbf{R}^{k\times 1}$ is the vector of unknown parameters, and the vector of random disturbances $u \in \mathbf{R}^{n\times 1}$ has a $N(0,\Sigma)$ distribution, where, as before, $diag(\Sigma) = (\sigma^2, ..., \sigma^2)$, such that the correlation matrix R corresponding to Σ has a maximum eigenvalue less than 2. Suppose, further, we want to test $H_0 : c'\beta = a$ against $H_A : c'\beta > a$ for some known vector $c \in \mathbf{R}^{k\times 1}$ on the basis of the $t-$statistic $T_c = (c'\beta - a)/\hat{sd}_{c'\hat{\beta}}$, where $\hat{sd}_{c'\hat{\beta}} = (\hat{\sigma}^2 c'(X'X)^{-1}c)^{1/2}$, $\hat{\sigma}^2 = (y - X\hat{\beta})'(y - X\hat{\beta})/(n-k)$. Using (24)-(26) we have

$$P(T_c > x) \le P(t_{n-k} > x) + (|R(2I_n - R)|^{-1/2} - 1)^{1/2}[P(t_{n-k} > x)]^{1/2},$$

$$P(T_c > x) \le |R(2I_n - R)|^{-1/4}[P(t_{n-k} > x)]^{1/2},$$

$$P(T_c > x) \le (e-1)P(t_{n-k} > x) - 0.5log(|\Sigma|/\sigma^{2n}),$$

where t_{n-k} denotes a r.v. with a $t-$distribution with $n-k$ degrees of freedom. A conservative critical region for the one-sided $t-$test with level α is given by $T_c > y_\alpha$, where y_α is such that

$$min\{P(t_{n-k} > y_\alpha) + (|R(2I_n - R)|^{-1/2} - 1)^{1/2}[P(t_{n-k} > y_\alpha)]^{1/2}, |R(2I_n - R)|^{-1/4}[P(t_{n-k} > y_\alpha)]^{1/2},$$

$$(e-1)P(t_{n-k} > y_\alpha) - 0.5log(|\Sigma|/\sigma^{2n}\} < \alpha.$$

Using the asymptotic properties of the statistics T_n and T_c, we conclude that the above critical regions can also be used in the case of non-Gaussian identically distributed errors when the sample size n is sufficiently large.

5. Proofs The proofs of Theorems 1-8 are based on the following reduction properties (Lemmas 1 and 2) for martingales and, more generally, multiplicative systems of an arbitrary order proved in Sharakhmetov and Ibragimov (2002) and de la Peña et. al. (2002).

Definition 1 *R.v.'s* $X_1, ..., X_n$ *form a multiplicative system of order* $\alpha \in \mathbf{N}$ *(shortly,* $MS(\alpha)$*) if* $E|X_j|^\alpha < \infty$, $j = 1, ..., n$, *and for any* $\alpha_j \in \{0, 1, ..., \alpha\}$, $j = 1, ..., n$,

$$E\prod_{j=1}^{n} X_j^{\alpha_j} = \prod_{j=1}^{n} EX_j^{\alpha_j}.$$

The systems $MS(1)$ and $MS(2)$ under the names multiplicative and strongly multiplicative systems, respectively, were introduced by Alexits (1961). Multiplicative systems of an arbitrary order were considered, e.g., by Kwapien (1987) and Sharakhmetov (1993). Examples of the multiplicative systems of order 1 $MS(1)$

are given, besides independent r.v.'s, by the lacunary trigonometric systems $\{cos2\pi n_k x, sin2\pi n_k x, k = 1, 2, ...\}$ on the interval $[0, 1]$ with the Lebesque measure for $n_{k+1}/n_k \geq 2$ and also by martingale-difference sequences. Examples of the systems $MS(2)$ are given by the lacunary trigonometric systems for $n_{k+1}/n_k \geq 3$ and martingale-difference sequences $X_1, ..., X_n$ with the non-random conditional variances $E(X_n^2|X_1, ..., X_{n-1}) = b_n^2 \in \mathbf{R}$, $n = 1, 2, ...$ Examples of the systems $MS(\alpha)$ include, for instance, the lacunary trigonometric systems with large lacunas, that is, with $n_{k+1}/n_k \geq \alpha + 1$ and also $\epsilon-$independent and asymptotically independent r.v.'s introduced by Zolotarev (1991) (see the discussion in de la Peña et. al. (2002)).

Lemma 1 *Let $\alpha \in \mathbf{N}$, and let A_i, $i = 1, ..., n$, be sets of real numbers such that $card(A_i) \leq \alpha + 1$, $i = 1, ..., n$. R.v.'s $X_1, ..., X_n$ taking values in $A_1, ..., A_n$, respectively, form a multiplicative system of order α if and only if they are jointly independent.*

Lemma 1 implies the following reduction property for general martingale-difference sequences.

Lemma 2 *A sequence of r.v.'s $\{X_n\}$ on a probability space $(\Omega, \Im, P)$ assuming two values is a martingale-difference with respect to an increasing sequence of $\sigma-$algebras $\Im_0 = (\Omega, \emptyset) \subseteq \Im_1 \subseteq ... \subseteq \Im$ if and only the r.v.'s $\{X_n\}$ are jointly independent.*

Since the r.v.'s $\eta_i = \epsilon_1^{r_{1i}} ... \epsilon_{i-1}^{r_{i-1,i}} \epsilon_i$, $i = 1, \ldots, n$, form a martingale-difference sequence with respect to the $\sigma-$algebras $\sigma(\epsilon_1, \epsilon_2, ..., \epsilon_i)$, $i = 1, ..., n$, we get the following corollary of Lemma 2 that describes the reduction properties of the summands in the random polynomials V_n in independent symmetric r.v.'s.

Lemma 3 *The r.v.'s $\eta_1, \ldots, \eta_n$ are jointly independent.*

Remark 1 *Let us note that Lemma 2 also implies Propositions 1 and 3 in Campbell and Dufour (1995). Let $w(z) = 1$ for $z \geq 0$, and $w(z) = 0$ for $z < 0$, and let $X_0, ..., X_{n-1}$ and $Y_1, ..., Y_n$ be r.v.'s such that Y_t is independent of $\sigma(X_0, X_1, ..., X_{t-1}, Y_1, ..., Y_{t-1})$ for each $t = 1, ..., n$, and $P(Y_t > 0) = P(Y_t < 0) = 1/2$ for $t = 1, ..., n$. Also, let $g_t = g_t(X_0, X_1, ..., X_t, Y_1, ..., Y_t)$, $t = 0, ..., n-1$, be a sequence of measurable functions of $X_0, ..., X_t$ such that $P(g_t = 0) = 0$ for $t = 0, ..., n-1$. According to Propositions 1 and 3 in Campbell and Dufour (1995), the statistic $S_g = \sum_{t=1}^n w(Y_t g_{t-1})$ has a binomial distribution $Bi(n, 0.5)$ with parameters n, 0.5, and, moreover, the r.v.'s $w(Y_t g_{t-1})$, $t = 1, ..., n$, are jointly independent if the r.v.'s $Y_1, ..., Y_n$ have continuous symmetric distributions. To see that the latter results follow from Lemma 2, it suffices to observe that the r.v.'s $2w(Y_t g_{t-1}) - 1$ form a martingale-difference sequence with respect to the $\sigma-$algebras $\sigma(X_0, X_1, ..., X_t, Y_1, ..., Y_t)$ under the assumptions of the propositions. In a recent paper, So and Shin (2001) considered sign tests for random walks against stationary alternative hypothesis in the model $y_t = h(x_t)$, $x_t = \rho(x_{t-1}, ..., x_{t-k}) + u_t$, $t = 1, ..., n$, where $\{y_t\}$, $t = 0, ..., n$, is a set of observations, $h(x_t)$ is an unknown monotone transformation of $\{x_t\}$, $\rho(x_k, ..., x_1)$ is an unknown regression function of interest, k is a positive integer, and $\{u_t\}$ is a sequence of errors satisfying the conditions*

A1: $\{sign(u_t)\}$ is a martingale difference sequence with respect to an increasing sequence of $\sigma-$fields $\{\Im_t\}$, $t = 1, ..., n$,

A2: $P(u_t = 0|\Im_{t-1}) = 0$,

where $sign(u_t)$ is the sign of u_t defined by $sign(u_t) = 1$ if $u_t > 0$, $sign(0) = 0$ and $sign(u_t) = -1$ otherwise. From the above, it follows that in fact the conditions A1 and A2 are equivalent to joint independence of the signs $sign(u_t)$. Moreover, using Lemma 2, one also immediately gets that the r.v.'s $sign(u_t)sign(v_{t-1})$, $t = 1, ..., n$, where v_t, $t = 1, ..., n$, is a sequence of $\Im_t-$measurable r.v.'s with no atom at zero, are jointly independent. This implies results in So and Shin (2001) concerning distributional properties of the test statistic based on the quantity $S_n(\rho) = \sum_{t=1}^n sign(u_t(\rho))sign(v_{t-1})$, where $u_t(\rho) = x_t - \rho(x_{t-1}, ..., x_{t-k})$.

Proof of Theorems 1-8. Let ϵ be a symmetric Bernoulli r.v. According to Hunt (1955), $Eg(Y) \le Eg(\epsilon)$ for all continuous convex functions $g : [-1, 1] \to \mathbf{R}$ and all r.v.'s Y such that $EY = 0$ and $|Y| \le 1$. The inequality implies that $Ef(cY + d) \le Ef(ca\epsilon + d)$ for all continuous convex functions $f : \mathbf{R} \to \mathbf{R}$, constants $a, c, d \in \mathbf{R}$, and all r.v.'s Y such that $EY = 0$ and $|Y| \le 1$. Using this fact and Lemma 3, we get that if $c_i \in \mathbf{R}$, $r_{ki} \in \{0, 1\}$, $k = 1, ..., i-1$, $i = 1, ..., n$, and $X_1, ..., X_n$ are independent r.v.'s such that $EX_i = 0$, $|X_i| \le 1$ (a.s.), $i = 1, ..., n$, then

$$Ef(\sum_{i=1}^n c_i X_1^{r_{1i}} ... X_{i-1}^{r_{i-1,i}} X_i) \le Ef(\sum_{i=1}^n c_i \epsilon_1^{r_{1i}} ... \epsilon_{i-1}^{r_{i-1,i}} \epsilon_i) = Ef(\sum_{i=1}^n c_i \epsilon_i) \tag{54}$$

for all continuous convex functions $f : \mathbf{R} \to \mathbf{R}$. From Corollary 2.5 in Pinelis (1994) and Theorem 1.1 in Figiel, Hitczenko, Johnson, Schechtman and Zinn (1997) it follows that

$$Ef(\sum_{i=1}^n c_i \epsilon_i) \le Ef(\sum_{i=1}^n c_i X_i') \tag{55}$$

for all twice differentiable even functions $f : \mathbf{R} \to \mathbf{R}$ such that f'' is convex and all independent symmetric r.v.'s X_i' such that $EX_i'^2 = 1$, $i = 1, ..., n$. Inequalities (54)-(55) imply (36). Inequality (37) follows letting the r.v.'s $\tilde{X}_i$ in (36) be the standard normal r.v.'s. Let us prove inequalities (12)-(16). Let $X_1, \ldots, X_n$ be independent r.v.'s such $EX_i = 0$, $|X_i| \le d_i$, $i = 1, \ldots, n$, and let $D^2 = \sum_{i=1}^n c_i^2 d_1^{2r_{1i}} \ldots d_2^{2r_{i-1,i}} d_i^2$. By Chebyshev's inequality we have

$$P(V_n > x) \le exp(-hx)Eexp(hV_n), \tag{56}$$

$x > 0$, $h > 0$. Using the above Hunt's inequality and Lemma 3, we get

$$Eexp(hV_n) \le Eexp(h \sum_{i=1}^n c_i d_1^{r_{1i}} ... d_{i-1}^{r_{i-1,i}} d_i \epsilon_i). \tag{57}$$

According to Hoeffding (1963),

$$Eexp(h \sum_{i=1}^n X_i) \le exp(\frac{1}{2} h^2 \sum_{i=1}^n d_i^2) \tag{58}$$

for all independent r.v.'s $X_1, \ldots, X_n$ such that $EX_i = 0$, $|X_i| \le d_i \in \mathbf{R}$. From (56)-(58) it follows that for $x > 0$

$$P(V_n > x) \le exp(\frac{1}{2}h^2 D^2 - hx). \tag{59}$$

The right-hand side of (59) has its minimum at $h = \frac{x}{D^2}$. Inserting this value in (59) we obtain inequality (12). From Chebyshev's inequality it follows that

$$P(V_n > x) = \frac{1}{2}P(|V_n| > x) \le \frac{1}{2}\frac{Ef(\frac{1}{D}|V_n|)}{f(\frac{x}{D})} \tag{60}$$

for all $f \in \overline{K}$ (introduced in Section 1). By (37),

$$Ef(\frac{1}{D}|V_n|) \le Ef(|Z|), \tag{61}$$

$f \in K$. Inequalities (60) and (61) imply (13). From the results obtained by Pinelis (1994) (see also Pinelis (1998)) it follows that inequality (61) implies that $P(V_n > x) \le \frac{2e^3}{9}(1 - \Phi(\frac{x}{D}))$, $x > 0$; $P(V_n > x) \le \frac{e^3}{9}\frac{\phi(\frac{x}{D})D}{x}$, $x > \sqrt{2}$, that is, (14) and (15) hold (note that the latter inequalities follow directly from Theorem 5.4 in Pinelis (1998) and the martingale structure of V_n). From (54), the fact that the function $f(x) = (|x| - u)_+^3$ belongs to $\overline{K}$ and the results obtained by Eaton (1970, 1974) (see also Dufour and Hallin (1993)) it follows that

$$E[(\frac{1}{D}|V_n| - u)_+]^3 \le E[(\frac{1}{D}|\sum_{i=1}^n c_i d_1^{r_{1i}}...d_{i-1}^{r_{i-1,i}} d_i\epsilon_i| - u)_+]^3 \le E[(|\frac{1}{\sqrt{n}}\sum_{i=1}^n \epsilon_i| - u)_+]^3. \tag{62}$$

Moreover,

$$E(\frac{1}{D}V_n)^2 \le \frac{1}{D^2}\sum_{i=1}^n c_i^2 d_1^{2r_{1i}}...d_{i-1}^{2r_{i-1,i}} d_i^2 E\epsilon_i^2 = 1. \tag{63}$$

Similarly to the proof of Proposition 1 in Dufour and Hallin (1993), relations (62) and (63) and Chebyshev's inequality give the first estimate in (16). The fact that $E[(\frac{1}{\sqrt{n}}|\sum_{i=1}^n \epsilon_i| - u)_+]^3 \le E[(|Z| - u)_+]^3$ by (9) and the first estimate in (16) imply the second inequality in (16). Since, according to Edelman (1990),

$$\inf_{0<u<x/D_j}\int_u^\infty ((t-u)^3/(\frac{x}{D_j} - u)^3)\phi(t)dt \le 1 - \Phi(x - \frac{1.5}{x}), \tag{64}$$

$x > 0$, we get the last estimate in (16).

Conditioning on $|X_1|, ..., |X_n|$ and using estimates (13)-(16) we obtain (similarly to Edelman (1986, 1990)) inequalities (17)-(21).

Let us prove Theorem 3. Using the relation

$$W_n = \sqrt{n}\frac{\overline{V}_n/s_n}{(1 - 1/n + \overline{V}_n^2/s_n^2)^{1/2}},$$

we obtain, similarly to Edelman (1986, 1990),

$$P(\sqrt{n}\,\overline{V}_n/s_n > x) = P(W_n > \frac{x}{(1 + (x^2-1)/n)^{1/2}}).$$

This and Theorem 2 imply the inequalities in Theorem 3.

The estimates in Theorems 4-6 follow from Theorems 1-3 and inequalities (24)-(26).

The expressions for the best constants in inequalities (38) follow from Lemma 3 and the results obtained by Haagerup (1982). The right-hand side inequality (38) and the estimate $(\sum_{i=1}^n z_i)^{t/2} \leq \sum_{i=1}^n n^{t/2-1} z_i^{t/2}$ for all $z_1, ..., z_n \geq 0$, $t \geq 2$, imply that estimate (39) holds with the constant $C^*(t, m)$ defined in Theorem 8. Sharpness of the constant $C^*(t, m)$ follows from the choice $c_i = 1/\sqrt{n}$, $i = 1, ..., n$, $X_i = \epsilon_i$, $i = 1, ..., n + h_m$, Lemma 3 and the central limit theorem.

Using the inequality $Ef(ca\epsilon + d) \leq Ef(cX + d)$ for all twice differentiable even functions $f : \mathbf{R} \to \mathbf{R}$ such that f'' is convex, constants $a, c, d \in \mathbf{R}$, and all symmetric r.v.'s X such that $EX^2 = a^2$ implied by Corollary 2.5 in Pinelis (1994) and Theorem 1.1 in Figiel, Hitczenko, Johnson, Schechtman and Zinn (1997), by induction and Lemma 3 we get

$$E|\sum_{i=1}^n c_i X_1^{r_{1i}} ... X_{i-1}^{r_{i-1,i}} X_i|^t \geq E|\sum_{i=1}^n c_i \epsilon_1^{r_{1i}} ... \epsilon_{i-1}^{r_{i-1,i}} \epsilon_i|^t = E|\sum_{i=1}^n \epsilon_i|^t \tag{65}$$

for independent identically distributed symmetric r.v.'s $X_1, ..., X_{n+h_m}$ with $EX_1^2 = 1$, $E|X_1|^t < \infty$, $t \geq 3$. Moreover, from estimate (3.28) in de la Peña et. al. (2003) and Lemma 3 it follows that

$$E|\sum_{i=1}^n c_i X_1^{r_{1i}} ... X_{i-1}^{r_{i-1,i}} X_i|^t \leq E|\sum_{i=1}^n \epsilon_i|^t + O(n) \tag{66}$$

for for independent identically distributed symmetric r.v.'s $X_1, ..., X_{n+h_m}$ with $EX_1^2 = 1$, $E|X_1|^t < \infty$, $2 < t \leq 4$. Relations (65) and (66) and the fact that, according to von Bahr (1965) and Michel (1976), $|E|\frac{1}{\sqrt{n}} \sum_{i=1}^n \epsilon_i|^t - E|Z|^t| = O(n^{1-t/2})$, $3 \leq t < 4$, imply relation (40).

REFERENCES

Adler, R. J., Feldman, R. E. and Gallagher, C. Analysing stable time series. In *A Practical Guide to Heavy Tails* (eds R. J. Adler, R. E. Feldman and M. S. Taqqu). Boston: Birkhauser, 113-58.

Alexits, G. (1961). *Convergence problems of orthogonal series.* International Series of Monographs in Pure and Applied Mathematics, Vol. 20, Pergamon Press, New York-Oxford-Paris.

Anderson, O. D. (1993). Exact general-lag serial correlation moments and approximate low-lag partial correlation moments for Gaussian white noise. *Journal of Time Series Analysis* **14**, 551-574.

Anderson, O. D. and Chen, Z.-G. (1996). Higher order moments of sample autocovariances and sample autocorrelations from an independent time series. *Journal of Time Series Analysis* **17**, 323-331.

Arcones, M. and Giné, E. (1993). Limit Theorems for $U-$Processes. *Annals of Probability* **21**, 1494-1542.

Bennett, G. (1962). Probability inequalities for the sum of independent random variables. *Journal of the American Statistical Association* **57**, 33-45.

Bernstein, S. (1946). *Theory of Probability.* Gostekhizdat, Moscow (in Russian).

Campbell, B. and Dufour, J.-M. (1995). Exact nonparametric orthogonality and random walk tests. *The Review of Economics and Statistics* **77**, 1-16

de la Peña, V. H. (1999). A general class of exponential inequalities for martingales and ratios. *Annals of Probability* **27**, 537-564.

de la Peña, V. H. and Giné, E. (1999). *Decoupling: from dependence to independence.* Probability and its Applications. Springer, New York, 392 pp.

de la Peña, V. H., Ibragimov, R. and Sharakhmetov, Sh. (2002). Characterizations of joint distributions, copulas, information, dependence and decoupling, with applications to time series. Working paper. Available at http://pantheon.yale.edu/~ri23/Papers/Copulas.pdf

de la Peña, V. H., Ibragimov, R. and Sharakhmetov, Sh. (2003). On extremal distributions and sharp L_p-bounds for sums of multilinear forms. *Annals of Probability* **31**, 630-675.

de la Peña, V. H., Klass, M. J. and Lai, T. L. (2000). Moment bounds for self-normalized processes. *High dimensional probability, II (Seattle, WA, 1999),* 3-12, *Progress in Probability,* Birkhäuser Boston, Boston, MA.

de la Peña, V. H., Klass, M. J. and Lai, T. L. (2002). Self-normalized processes: exponential inequalities, moment bounds and iterated logarithm laws. Working Paper.

de la Peña, V. H. and Montgomery-Smith, S. (1995). Decoupling inequalities for the tail probabilities of multivariate $U-$statistics. *Annals of Probability* **23**, 806-816.

de la Peña, V. H. and Zamfirescu, I.-M. (2002). Decoupling and domination inequalities with application to Wald's identity for martingales. *Statistics and Probability Letters* **57**, 157-170.

Dharmadhikari, S. W. and Jogdeo, K. (1969). Bounds on moments of certain random variables. *Annals of Mathematical Statistics* **40**, 1506-1509.

Dufour, J.-M. and Hallin, M. (1991). Nonuniform bounds for nonparametric t-tests. *Econometric Theory* **7**, 253-263.

Dufour, J.-M. and Hallin, M. (1992). Simple exact bounds for distributions of linear signed rank statistics. *Journal of Statistical Planning and Inference* **31**, 311-333.

Dufour, J.-M. and Hallin, M. (1993). Improved Eaton bounds for linear combinations of bounded random variables, with statistical applications. *Journal of the American Statistical Association* **88**, 1026-1033.

Eaton, M. L. (1970). A note on symmetric Bernoulli random variables. *Annals of Mathematical Statistics* **41**, 1223-1226.

Eaton, M. L. (1974). A probability inequality for linear combinations of bounded random variables. *Annals of Statistics* **2**, 609-614.

Eaton, M. L. and Efron, B. (1970). Hotelling's T^2 test under symmetry conditions. *Journal of the American Statistical Association* **65**, 702-711.

Edelman, D. (1986). Bounds for a nonparametric t table. *Biometrika* **73**, 242-243.

Edelman, D. (1990). An inequality of optimal order for the tail probabilities of the t-statistic under symmetry. *Journal of the American Statistical Association* **85**, 120-122.

Efron, B. (1969). Student's t-test under symmetry conditions. *Journal of the American Statistical Association* **64**, 1278-1302.

Eichelsbacher, P. (2000). Moderate deviations for degenerate U-processes. *Stochastic Processes and their Applications* **87**, 255-279

Embrechts, P., Kluppelberg, C. and Mikosch, T. (1997). *Modelling extremal events for insurance and finance.* Springer, 645 pp.

Figiel, T., Hitczenko, P., Johnson, W. B., Schechtman, G. and Zinn, J. (1997). Extremal properties of Rademacher functions with applications to the Khintchine and Rosenthal inequalities. *Transactions of the American Mathematical Society* **349**, 997-1027.

Giné, E., Latała, R. and Zinn, J. (2000). Exponential and moment inequalities for $U-$statistics. *High dimensional probability, II (Seattle, WA, 1999),* 13-38, *Progress in Probability,* Birkhäuser Boston, Boston, MA.

Haagerup, U. (1982). The best constants in the Khintchine inequality. *Studia Mathematica* **70**, 231-283.

Heilig, C. and Nolan, D. (2001). Limit theorems for the infinite-degree $U-$statistics. *Statistica Sinica* **11**, 289-302.

Ho, H.-C. and Hsing, T. (1997). Limit theorems for functionals of moving averages. *Annals of Probability* **25**, 1636-1669.

Hoeffding, W. (1963). Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association* **58**, 13-30.

Hunt, G. A. (1955). An inequality in probability theory. *Proceedings of the American Mathematical Society* **6**, 506-510.

Joe, H. (1987). Majorization, randomness and dependence for multivariate distributions. *Annals of Probability* **15**, 1217-1225.

Joe, H. (1989). Relative entropy measures of multivariate dependence. *Proceedings of the American Mathematical Society* **84**, 157-164.

Jurečkovà, J. and Sen, P. K. (1996). *Robust statistical procedures. Asymptotics and interrelations.* Wiley Series in Probability and Statistics: Applied Probability and Statistics. A Wiley-Interscience Publication. John Wiley & Sons, New York.

Kariya, T. and Sinha, B. K. (1989). *Robustness of statistical tests.* Statistical Modeling and Decision Science. Academic Press, Boston, MA.

Krakowiak, W. and Szulga, J. (1986). Random multilinear forms. *Annals of Probability* **14**, 955-973.

Kwapień, S. (1987). Decoupling inequalities for polynomial chaos. *Annals of Probability* **15**, 1062-1071.

Kwapien, S. and Woyczynski, W. (1992). *Random series and stochastic integrals: single and multiple.* Birkhauser, Boston.

Loretan, M. and Phillips, P. C. B. (1994). Testing the covariance stationarity of heavy-tailed time series. *Journal of Empirical Finance* **3**, 211-248.

McConnell, T. R. and Taqqu, M. (1986). Decoupling inequalities for multilinear forms in independent symmetric random variables. *Annals of Probability* **14**, 943-954.

Michel, R. (1976). Nonuniform central limit bounds with applications to probabilities of deviations. *Annals of Probability* **4**, 102-106.

Phillips, P. C. B. and Solo, V. (1992). Asymptotics for linear processes. *Annals of Statistics* **20**, 971-1001.

Pinelis, I. (1994). Extremal probabilistic problems and Hotelling's T^2 test under a symmetry condition. *Annals of Statistics* **22**, 357-368.

Pinelis, I. (1998). Optimal tail comparison based on comparison of moments. In *High dimensional probability (Oberwolfach, 1996),* 297-314, *Progress in Probability* **43**, Birkhäuser, Basel.

Prokhorov, Yu. V. (1959). An extremal problem in probability theory. *Theory of Probability and Its Applications* **4**, 201-203.

Sharakhmetov, Sh. (1993). r-independent random variables and multiplicative systems. (Russian) Dopov./ Dokl. Akad. Nauk Ukraïni, 43-45.

Sharakhmetov, Sh. and Ibragimov, R. (2002). A characterization of joint distribution of two-valued random variables and its applications. *Journal of Multivariate Analysis* **83**, 389-408.

So, B. S. and Shin, D. W. (2001). An invariant sign test for random walks based on recursive median adjustment. *Journal of Econometrics* **102**, 197-229.

Szulga, J. (1998). *Introduction to random chaos.* Chapman and Hall, 297 pp.

Talagrand, M. (1996). A new look at independence. *Annals of Probability* **24**, 1-34.

von Bahr, B. (1965). On the convergence of moments in the central limit theorem. *Annals of Mathematical Statistics* **36**, 808-818.

Tse, Y. K. and Zhang, X. B. (2002). The variance ratio test with stable Paretian errors. *Journal of Time Series Analysis* **23**, 117-126.

Zolotarev, V. M. (1991). Reflection on the classical theory of limit theorems. I. *Theory of Probability and Its Applications* **36**, 124-137.

Table 1

Empirical sizes (%) of the two-sided conservative sign tests based on bounds (42) and (43) for model (41)													
Error distribution	D1, Normal			D2, Variance Mixture of Normals			D3, Normal with heteroskedasticity		D4, Cauchy		D5, Normal-Cauchy	
Sample size	Bound (42)	Bound (43)	Limit, autoregression	Bound (42)	Bound (43)	Limit, autoregression	Bound (42)	Bound (43)	Bound (42)	Bound (43)	Bound (42)	Bound (43)
20	0.4	0.5	4.3	0.4	0.2	4.5	0.1	0.1	0.1	0.1	0.1	0.1
50	0.7	0.8	4.9	0.7	0.6	4.8	0.6	0.6	0.1	0.1	0.1	0.1
100	0.9	0.9	4.9	0.8	0.9	5.0	0.8	0.8	0.1	0.1	0.1	0.1
200	0.8	0.8	5.0	0.8	1.0	4.8	1.2	0.9	0.1	0.2	0.1	0.1
500	1.2	0.9	5.1	1.0	1.1	4.9	1.0	0.9	0.2	0.1	0.1	0.1

Nominal test size: 0.05; the number of replications: 10,000

Table 2

Empirical sizes (%) of the two-sided conservative sign tests based on bounds (42) and (43) for model (41)												
Error distribution	D1, Normal			D2, Variance Mixture of Normals			D3, Normal with heteroskedasticity		D4, Cauchy		D5, Normal-Cauchy	
Sample size	Bound (42)	Bound (43)	Limit, autoregression	Bound (42)	Bound (43)	Limit, autoregression	Bound (42)	Bound (43)	Bound (42)	Bound (43)	Bound (42)	Bound (43)
20	0.4	0.3	4.3	0.5	0.4	4.5	0.3	0.4	0.4	0.6	0.4	0.2
50	0.9	0.8	4.9	0.8	0.9	4.8	0.9	1.1	0.9	0.8	0.9	1.0
100	0.8	0.7	4.9	0.9	0.8	5.0	0.7	0.8	0.9	0.9	0.8	0.9
200	1.0	1.0	5.0	1.1	1.0	4.8	1.0	1.1	1.1	1.1	1.1	1.0
500	1.2	1.2	5.1	1.1	1.0	4.9	1.2	1.2	1.1	1.3	1.2	1.1

Nominal test size: 0.05; the number of replications: 10,000

Table 3a

Empirical powers (%) of the two-sided conservative tests based on bounds (42) and (43) for model (41) under ρ=0.5													
Error distribution	D1, Normal			D2, Variance Mixture of Normals			D3, Normal with heteroskedasticity		D4, Cauchy		D5, Normal-Cauchy	
Sample size	Bound (42)	Bound (43)	Limit, autoregression	Bound (42)	Bound (43)	Limit, autoregression	Bound (42)	Bound (43)	Bound (42)	Bound (43)	Bound (42)	Bound (43)
20	14.2	14.5	56.6	14.4	14.3	57.9	1.8	2.3	2.4	2.8	2.9	3.1
50	73.3	73.5	93.4	74.0	74.2	93.5	33.8	33.2	11.8	12.1	10.2	10.7
100	98.9	98.7	99.9	98.9	98.7	99.9	81.6	83.1	18.1	18.8	15.3	16.1
200	100.0	100.0	100.0	100.0	100.0	100.0	99.4	99.6	22.1	23.0	20.6	20.3
500	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	24.3	24.9	24.2	24.6

Nominal test size: 0.05; the number of replications: 10,000

Table 3b

Empirical powers (%) of the two-sided conservative tests based on bounds (42) and (43) for model (41) under ρ=–0.5												
Error distribution	D1, Normal			D2, Variance Mixture of Normals			D3, Normal with heteroskedasticity		D4, Cauchy		D5, Normal-Cauchy	
Sample size	Bound (42)	Bound (43)	Limit, autoregression	Bound (42)	Bound (43)	Limit, autoregression	Bound (42)	Bound (43)	Bound (42)	Bound (43)	Bound (42)	Bound (43)
20	13.7	14.9	57.5	14.0	14.1	57.6	2.0	1.9	2.9	3.0	2.8	3.0
50	73.2	74.1	93.1	73.5	73.5	93.7	33.0	34.1	11.3	11.4	10.5	10.3
100	98.8	98.8	99.8	98.6	98.7	99.8	81.5	82.9	18.1	18.7	15.8	15.7
200	100.0	100.0	100.0	100.0	100.0	100.0	99.5	99.6	23.1	23.1	20.7	21.2
500	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	23.8	25.1	24.0	24.6

Nominal test size: 0.05; the number of replications: 10,000

Table 4a

Empirical powers (%) of the two-sided conservative sign tests based on bounds (42) and (43) for model (41) under ρ=0.5													
Error distribution	D1, Normal			D2, Variance Mixture of Normals			D3, Normal with heteroskedasticity		D4, Cauchy		D5, Normal-Cauchy	
Sample size	Bound (42)	Bound (43)	Limit, autoregression	Bound (42)	Bound (43)	Limit, autoregression	Bound (42)	Bound (43)	Bound (42)	Bound (43)	Bound (42)	Bound (43)
20	9.3	9.1	56.6	9.3	9.0	57.9	8.1	7.8	26.2	25.3	14.5	14.1
50	41.8	41.2	93.4	41.2	41.9	93.5	39.3	39.4	83.8	82.6	61.7	62.7
100	76.3	76.5	99.9	74.9	75.9	99.9	75.7	74.7	99.3	99.3	93.3	93.8
200	98.3	98.2	100.0	98.0	98.2	100.0	98.6	98.1	100.0	100.0	100.0	100.0
500	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Nominal test size: 0.05; the number of replications: 10,000

Table 4b

Empirical powers (%) of the two-sided conservative tests based on bounds (42) and (43) for model (41) under ρ=–0.5												
Error distribution	D1, Normal			D2, Variance Mixture of Normals			D3, Normal with heteroskedasticity		D4, Cauchy		D5, Normal-Cauchy	
Sample size	Bound (42)	Bound (43)	Limit, autoregression	Bound (42)	Bound (43)	Limit, autoregression	Bound (42)	Bound (43)	Bound (42)	Bound (43)	Bound (42)	Bound (43)
20	9.3	9.0	57.5	8.4	9.0	57.6	7.9	7.9	26.5	26.1	14.6	14.8
50	41.2	41.2	93.1	41.2	41.4	93.7	39.5	39.3	83.4	83.4	63.4	61.0
100	76.4	75.5	99.8	75.1	76.0	99.8	74.9	75.6	99.2	99.1	93.8	93.7
200	98.1	98.4	100.0	98.3	98.2	100.0	98.2	98.0	100.0	100.0	99.9	99.9
500	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Nominal test size: 0.05; the number of replications: 10,000