The testcorr Package

Violetta Dalla
National and Kapodistrian University of Athens

Liudas Giraitis
Queen Mary University of London

Peter C. B. Phillips
Yale University, University of Auckland, University of Southampton, Singapore Management University

Abstract

The R package **testcorr** implements standard and robust procedures for testing the significance of the autocorrelation in univariate data and the cross-correlation in bivariate data. It also includes tests for the significance of pairwise Pearson correlation in multivariate data and the i.i.d. property for univariate data. The standard testing procedures on significance of correlation are used commonly by practitioners while their robust versions were developed in Dalla, Giraitis, and Phillips (2020), where the tests for i.i.d. property can be also found. This document briefly outlines the testing procedures and provides simple examples.

Keywords: autocorrelation, cross-correlation, Pearson correlation, i.i.d., R.

1. Introduction

Inference on the significance of the autocorrelation $\rho_k = \text{corr}(x_t, x_{t-k})$ or the cross-correlation $\rho_{xy,k} = \text{corr}(x_t, y_{t-k})$ is a common first step in the analysis of univariate $\{x_t\}$ or bivariate $\{x_t, y_t\}$ time series data. Moreover, it is common to test the significance of pair-wise correlations $\rho_{x_i x_j} = \text{corr}(x_{it}, x_{jt})$ in multivariate $\{x_{1t}, x_{2t}, ..., x_{pt}\}$ data, cross-sectional or time series. Standard inference procedures[1] are valid for i.i.d. univariate or mutually independent bivariate/multivariate data and their size can be significantly distorted otherwise, in particular, by heteroscedasticity and dependence. The robust methods[2] given in Dalla *et al.* (2020) allow testing for significant autocorrelation/cross-correlation/correlation under more general settings, e.g., they allow for heteroscedasticity and dependence in each series and mutual dependence across series.

The R (R Core Team 2019) package **testcorr** includes the functions `ac.test` and `cc.test` that implement the standard and robust procedures for testing significance of autocorrelation and cross-correlation, respectively. Moreover, the package provides the function `rcorr.test` that evaluates the sample Pearson correlation matrix for multivariate data with robust p-values for testing sig-

[1] Like those implemented in the **stats** (R Core Team 2019), **sarima** (Boshnakov and Halliday 2019), **portes** (Mahdi and McLeod 2018) and **Hmisc** (Harrell Jr, with contributions from Charles Dupont *et al.* 2019) packages, functions `stats::acf`, `stats::ccf`, `stats::Box.test`, `sarima::acfIidTest`, `sarima::whiteNoiseTest` with `h0 = "iid"`, `portes::LjungBox`, `stats::cor.test` and `Hmisc::rcorr`.

[2] These robust methods are valid under more general settings compared to those in the **sarima** (Boshnakov and Halliday 2019) and **normwhn.test** (Wickham 2012) packages, functions `sarima::acfGarchTest`, `sarima::acfWnTest`, `sarima::whiteNoiseTest` with `h0 = "garch"` and `normwhn.test::whitenoise.test`.

nificance of its elements. The package also contains the function `iid.test` that conducts testing procedures for the i.i.d. property[3] of univariate data introduced in Dalla *et al.* (2020). Sections 2-5 describe the testing procedures that each function implements and provide examples. Section 6 outlines some suggestions relating to the application of the testing procedures.

2. Testing zero autocorrelation: `ac.test`

For a univariate time series $\{x_t\}$, given a sample $x_1, ..., x_n$, the null hypothesis $H_0 : \rho_k = 0$ of no autocorrelation at lag $k = 1, 2, ...$ is tested at α significance level using the sample autocorrelation $\widehat{\rho}_k$ and the $100(1-\alpha)\%$ confidence band (CB) for zero autocorrelation, obtained using the corresponding t-type statistics (t_k "standard" and $\widetilde{t}_k$ "robust").[4,5] The null hypothesis $H_0 : \rho_1 = ... = \rho_m = 0$ of no autocorrelation at cumulative lags $m = 1, 2, ...$ is tested using portmanteau type statistics (Ljung-Box LB_m "standard" and $\widetilde{Q}_m$ "robust").[6] The following notation is used.

Standard procedures:

$$CB(100(1-\alpha)\%) = (-z_{\alpha/2}/\sqrt{n}, z_{\alpha/2}/\sqrt{n}), \quad t_k = \sqrt{n}\widehat{\rho}_k, \quad LB_m = (n+2)n\sum_{k=1}^{m}\frac{\widehat{\rho}_k^2}{n-k}.$$

Robust procedures:

$$CB(100(1-\alpha)\%) = (-z_{\alpha/2}\frac{\widehat{\rho}_k}{\widetilde{t}_k}, z_{\alpha/2}\frac{\widehat{\rho}_k}{\widetilde{t}_k}), \quad \widetilde{t}_k = \frac{\sum_{t=k+1}^{n} e_{tk}}{(\sum_{t=k+1}^{n} e_{tk}^2)^{1/2}}, \quad \widetilde{Q}_m = \widetilde{t}'\,\widehat{R}^{*\,-1}\,\widetilde{t},$$

where $e_{tk} = (x_t - \bar{x})(x_{t-k} - \bar{x})$, $\bar{x} = n^{-1}\sum_{t=1}^{n} x_t$, $\widetilde{t} = (\widetilde{t}_1, ..., \widetilde{t}_m)'$ and $\widehat{R}^* = (\widehat{r}_{jk}^*)$ is a matrix with elements $\widehat{r}_{jk}^* = \widehat{r}_{jk} I(|\tau_{jk}| > \lambda)$ where λ is the threshold,

$$\widehat{r}_{jk} = \frac{\sum_{t=\max(j,k)+1}^{n} e_{tj}e_{tk}}{(\sum_{t=\max(j,k)+1}^{n} e_{tj}^2)^{1/2}(\sum_{t=\max(j,k)+1}^{n} e_{tk}^2)^{1/2}}, \quad \tau_{jk} = \frac{\sum_{t=\max(j,k)+1}^{n} e_{tj}e_{tk}}{(\sum_{t=\max(j,k)+1}^{n} e_{tj}^2 e_{tk}^2)^{1/2}}.$$

Applying standard and robust tests, at significance level α, $H_0 : \rho_k = 0$ is rejected when $\widehat{\rho}_k \notin CB(100(1-\alpha)\%)$ or $|t_k|, |\widetilde{t}_k| > z_{\alpha/2}$. In turn, $H_0 : \rho_1 = ... = \rho_m = 0$ is rejected when $LB_m, \widetilde{Q}_m > \chi^2_{m,\alpha}$. Here, $z_{\alpha/2}$ and $\chi^2_{m,\alpha}$ stand for the upper $\alpha/2$ and α quantiles of $N(0,1)$ and χ^2_m distributions.

Example

We provide an example to illustrate testing for zero autocorrelation of a univariate time series $\{x_t\}$ using the function `ac.test`. We simulate $n = 300$ data as GARCH(1,1): $x_t = \sigma_t \varepsilon_t$ with $\sigma_t^2 = 1 + 0.2x_{t-1}^2 + 0.7\sigma_{t-1}^2$ and $\varepsilon_t \sim$ i.i.d. $N(0,1)$.[7] The series $\{x_t\}$ is not autocorrelated but is not i.i.d. This is one of the models examined in the Monte Carlo study of Dalla *et al.* (2020). They

[3]Existing procedures include the rank test, the turning point test, the test for a Bernoulli scheme and the difference-sign test which are included in the package **spgs** (Hart and Martínez 2018), functions `spgs::rank.test`, `spgs::turningpoint.test`, `spgs::diid.test`, `spgs::diffsign.test`.

[4]Robust CB for zero autocorrelation provides a robust acceptance region for H_0.

[5]The standard procedure is implemented by `stats::acf`, `sarima::acfIidTest` and `sarima::whiteNoiseTest` with `h0 = "iid"`.

[6]The standard procedure is implemented by `stats::Box.test`, `sarima::acfIidTest`, `sarima::whiteNoiseTest` with `h0 = "iid"` and `portes::LjungBox`.

[7]We initialize $\sigma_1^2 = \text{var}(x_t) = 10$, simulate 400 observations and drop the first 100.

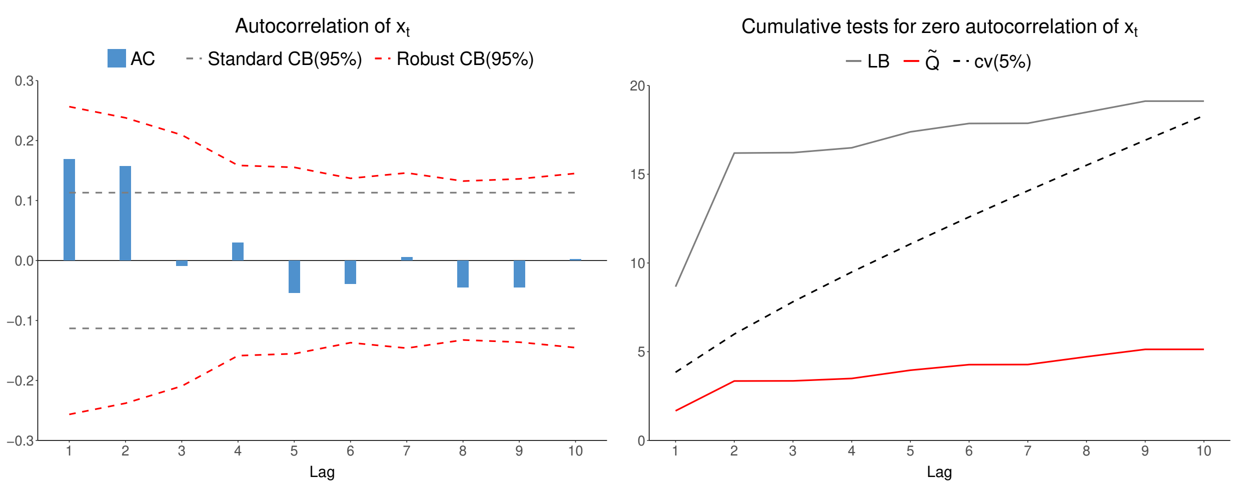
find that the standard testing procedures are a bit oversized (e.g. by around 8% when $k, m = 1$), while the robust tests are correctly sized. We choose a realization where this is evident.

```
R> set.seed(1798)
R> e <- rnorm(400)
R> x <- matrix(0, nrow = 400, ncol = 1)
R> s2 <- matrix(0, nrow = 400, ncol = 1)
R> s2[1] <- 10
R> x[1] <- sqrt(s2[1]) * e[1]
R> for (t in 2:400) {
R>   s2[t] <- 1 + 0.2 * (x[t - 1] ^ 2) + 0.7 * s2[t - 1]
R>   x[t] <- sqrt(s2[t]) * e[t]
R> }
R> x <- x[101:400]
```

We use the function `ac.test` to evaluate the results on testing for maximum 10 lags at significance level $\alpha = 5\%$ with threshold $\lambda = 2.576$. The plots are shown in the Plots pane and the table is printed on the Console. We don't pass any variable's name and set to 2 the scaling factor of the fonts in the plots.[8]

```
R> library(testcorr)
R> ac.test(x, max.lag = 10, alpha = 0.05, lambda = 2.576, plot = TRUE,
+          table = TRUE, var.name = NULL, scale.font = 2)
```

We have the following testing outputs:



[8]The default values are `alpha = 0.05`, `lambda = 2.576`, `plot = TRUE`, `table = TRUE`, `var.name = NULL` and `scale.font = 1`. Setting `scale.font = 2` is useful in order to upscale the fonts in the plots in order to export them as displayed here; the default value is suggested for viewing the plots in the Plots pane.

Tests for zero autocorrelation of x

Lag	AC	Stand. CB(95%)	Robust CB(95%)	Lag	t	p-value	t-tilde	p-value	Lag	LB	p-value	Q-tilde	p-value
1	0.169	(-0.113, 0.113)	(-0.257, 0.257)	1	2.929	0.003	1.292	0.196	1	8.664	0.003	1.669	0.196
2	0.157	(-0.113, 0.113)	(-0.238, 0.238)	2	2.726	0.006	1.296	0.195	2	16.194	0.000	3.348	0.187
3	-0.009	(-0.113, 0.113)	(-0.209, 0.209)	3	-0.153	0.878	-0.083	0.934	3	16.218	0.001	3.355	0.340
4	0.030	(-0.113, 0.113)	(-0.159, 0.159)	4	0.517	0.605	0.369	0.712	4	16.491	0.002	3.491	0.479
5	-0.054	(-0.113, 0.113)	(-0.155, 0.155)	5	-0.937	0.349	-0.682	0.495	5	17.390	0.004	3.957	0.556
6	-0.039	(-0.113, 0.113)	(-0.137, 0.137)	6	-0.678	0.498	-0.560	0.576	6	17.862	0.007	4.270	0.640
7	0.006	(-0.113, 0.113)	(-0.146, 0.146)	7	0.101	0.920	0.078	0.938	7	17.872	0.013	4.276	0.747
8	-0.045	(-0.113, 0.113)	(-0.132, 0.132)	8	-0.777	0.437	-0.664	0.507	8	18.497	0.018	4.717	0.787
9	-0.045	(-0.113, 0.113)	(-0.136, 0.136)	9	-0.775	0.438	-0.645	0.519	9	19.121	0.024	5.132	0.823
10	0.002	(-0.113, 0.113)	(-0.145, 0.145)	10	0.036	0.972	0.028	0.978	10	19.122	0.039	5.133	0.882

The left-hand side plot is graphing for maximum 10 lags, the sample autocorrelation $\widehat{\rho}_k$ (`"AC"`), the standard and robust CB(95%). The right-hand side plot is graphing for maximum 10 lags, the cumulative test statistics LB_m, $\widetilde{Q}_m$ and their critical values at 5% significance level (`"cv(5%)"`). The table reports the results of the plots along with the p-values for all the statistics: standard t_k (`"t"`) and LB_m (`"LB"`) and robust $\widetilde{t}_k$ (`"t-tilde"`) and $\widetilde{Q}_m$ (`"Q-tilde"`). The columns of the table can each be extracted by adding `$lag`, `$ac`, `$scb`, `$rcb`, `$t`, `$pvt`, `$ttilde`, `$pvttilde`, `$lb`, `$pvlb`, `$qtilde`, `$pvqtilde` at the end of the function call.

From the left-hand side plot we can conclude that $H_0 : \rho_k = 0$ is rejected at $\alpha = 5\%$ when $k = 1, 2$ and is not rejected at $\alpha = 5\%$ when $k = 3, ..., 10$ using standard methods, but is not rejected at $\alpha = 5\%$ for any k using robust methods. From the right-hand side plot we can conclude that the cumulative hypothesis $H_0 : \rho_1 = ... = \rho_m = 0$ is rejected at $\alpha = 5\%$ for all m using standard methods, but is not rejected at any m using robust methods. Subsequently, from the p-values in the table we find that using standard methods, $H_0 : \rho_k = 0$ is rejected at $\alpha = 1\%$ when $k = 1, 2$ and is not rejected at $\alpha = 10\%$ when $k = 3, ..., 10$, whereas using robust methods it is not rejected at $\alpha = 10\%$ for any k. Using standard methods the cumulative hypothesis $H_0 : \rho_1 = ... = \rho_m = 0$ is rejected at $\alpha = 0.1\%$ for $m = 2$, at $\alpha = 1\%$ when $m = 1, 3, ..., 6$ and at $\alpha = 5\%$ for $m = 7, ..., 10$, whereas using robust methods it is not rejected at $\alpha = 10\%$ for any m. Overall, standard testing procedures show evidence of autocorrelation, although the series is not autocorrelated. The robust testing procedures provide the correct inference.

3. Testing zero cross-correlation: `cc.test`

For a bivariate time series $\{x_t, y_t\}$, given a sample $(x_1, ..., x_n), (y_1, ..., y_n)$, the null hypothesis $H_0 : \rho_{xy,k} = 0$ of no cross-correlation at lag $k = 0, 1, 2, ...$ is tested at α significance level using the sample cross-correlation $\widehat{\rho}_{xy,k}$ and the $100(1-\alpha)\%$ confidence band (CB) for zero cross-correlation, obtained using the corresponding t-type statistics ($t_{xy,k}$ "standard" and $\widetilde{t}_{xy,k}$ "robust").[9,10] The null hypothesis $H_0 : \rho_{xy,0} = ... = \rho_{xy,m} = 0$ of no cross-correlation at cumulative lags $m = 0, 1, 2, ...$ is tested using portmanteau type statistics (Haugh-Box $HB_{xy,m}$ "standard" and $\widetilde{Q}_{xy,m}$ "robust").[11] The following notation is used.

Standard procedures:

$$CB(100(1-\alpha)\%) = (-z_{\alpha/2}/\sqrt{n}, z_{\alpha/2}/\sqrt{n}), \quad t_{xy,k} = \sqrt{n}\widehat{\rho}_{xy,k}, \quad HB_{xy,m} = n^2 \sum_{k=0}^{m} \frac{\widehat{\rho}^2_{xy,k}}{n-k}.$$

[9] Robust CB for zero cross-correlation provides a robust acceptance region for H_0.
[10] The standard procedure is implemented by `stats::ccf`.
[11] The standard procedure is not provided in any R package. A version of the Haugh-Box statistic involving also the autocorrelations of each series is implemented by `portes::LjungBox`.

Robust procedures:

$$CB(100(1-\alpha)\%) = (-z_{\alpha/2}\frac{\widehat{\rho}_{xy,k}}{\widetilde{t}_{xy,k}}, z_{\alpha/2}\frac{\widehat{\rho}_{xy,k}}{\widetilde{t}_{xy,k}}), \quad \widetilde{t}_{xy,k} = \frac{\sum_{t=k+1}^{n} e_{xy,tk}}{\left(\sum_{t=k+1}^{n} e_{xy,tk}^2\right)^{1/2}}, \quad \widetilde{Q}_{xy,m} = \widetilde{t}_{xy}^{\prime}\, \widehat{R}_{xy}^{*\,-1}\, \widetilde{t}_{xy},$$

where $e_{xy,tk} = (x_t - \bar{x})(y_{t-k} - \bar{y})$, $\bar{x} = n^{-1}\sum_{t=1}^{n} x_t$, $\bar{y} = n^{-1}\sum_{t=1}^{n} y_t$, $\widetilde{t}_{xy} = (\widetilde{t}_{xy,0}, ..., \widetilde{t}_{xy,m})'$ and $\widehat{R}_{xy}^{*} = (\widehat{r}_{xy,jk}^{*})$ is a matrix with elements $\widehat{r}_{xy,jk}^{*} = \widehat{r}_{xy,jk} I(|\tau_{xy,jk}| > \lambda)$ where λ is the threshold,

$$\widehat{r}_{xy,jk} = \frac{\sum_{t=\max(j,k)+1}^{n} e_{xy,tj} e_{xy,tk}}{(\sum_{t=\max(j,k)+1}^{n} e_{xy,tj}^2)^{1/2} (\sum_{t=\max(j,k)+1}^{n} e_{xy,tk}^2)^{1/2}}, \quad \tau_{xy,jk} = \frac{\sum_{t=\max(j,k)+1}^{n} e_{xy,tj} e_{xy,tk}}{(\sum_{t=\max(j,k)+1}^{n} e_{xy,tj}^2 e_{xy,tk}^2)^{1/2}}.$$

Applying standard and robust tests, at significance level α, $H_0 : \rho_{xy,k} = 0$ is rejected when $\widehat{\rho}_{xy,k} \notin CB(100(1-\alpha)\%)$ or $|t_{xy,k}|, |\widetilde{t}_{xy,k}| > z_{\alpha/2}$. In turn, $H_0 : \rho_{xy,0} = ... = \rho_{xy,m} = 0$ is rejected when $HB_{xy,m}, \widetilde{Q}_{xy,m} > \chi^2_{m,\alpha}$. Here, $z_{\alpha/2}$ and $\chi^2_{m,\alpha}$ stand for the upper $\alpha/2$ and α quantiles of $N(0,1)$ and χ^2_m distributions.

The above procedures where outlined for $k, m \geq 0$. For $k, m < 0$, the tests are analogously defined, noting that $\widehat{\rho}_{xy,k} = \widehat{\rho}_{yx,-k}$, $t_{xy,k} = t_{yx,-k}$, $\widetilde{t}_{xy,k} = \widetilde{t}_{yx,-k}$, $HB_{xy,m} = HB_{yx,-m}$, $\widetilde{Q}_{xy,m} = \widetilde{Q}_{yx,-m}$.

Example

We provide an example to illustrate testing for zero cross-correlation of a bivariate time series $\{x_t, y_t\}$ using the function `cc.test`. We simulate $n = 300$ data as noise and SV-AR(1) using the same noise in the AR(1) part: $x_t = \varepsilon_t$ and $y_t = \exp(z_t)u_t$ with $z_t = 0.7z_{t-1} + \varepsilon_t$, $\varepsilon_t, u_t \sim$ i.i.d. $N(0,1)$, $\{\varepsilon_t\}$ and $\{u_t\}$ mutually independent.[12] The series $\{x_t\}$ and $\{y_t\}$ are uncorrelated but are not independent of each other, both are serially uncorrelated and only $\{x_t\}$ is i.i.d. This is one of the models examined in the Monte Carlo study of Dalla *et al.* (2020). They find that the standard testing procedures are rather oversized (e.g. by around 25% when $k, m = 0$), while the robust tests are correctly sized. We choose a realization where this is evident.

```
R> set.seed(227)
R> e <- rnorm(400)
R> set.seed(492)
R> u <- rnorm(300)
R> x <- e[101:400]
R> z <- matrix(0, nrow = 400, ncol = 1)
R> for (t in 2:400) {
R>   z[t] <- 0.7 * z[t - 1] + e[t]
R> }
R> z <- z[101:400]
R> y <- exp(z) * u
```

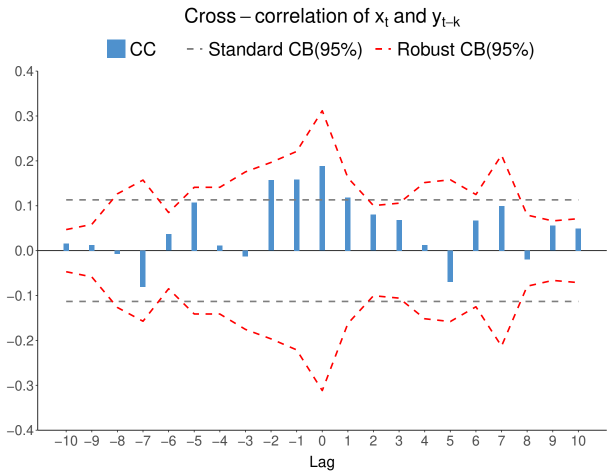
We use the function `cc.test` to evaluate the results on testing for maximum ±10 lags at significance level $\alpha = 5\%$ with threshold $\lambda = 2.576$. The plots are shown in the Plots pane and the table is

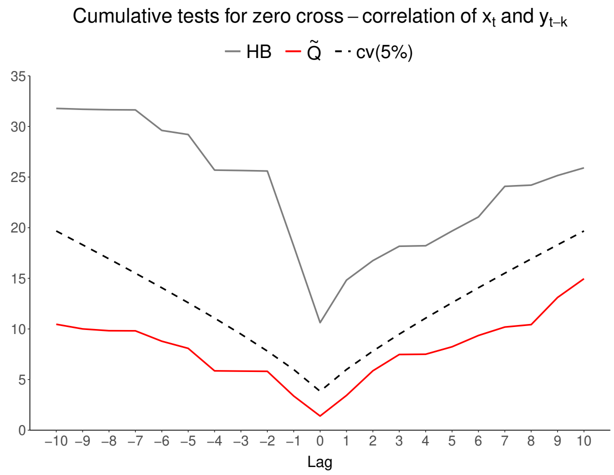
[12] We initialize $z_1 = Ez_t = 0$, simulate 400 observations and drop the first 100.

printed on the Console. We don't pass any variables' names and set to 2 the scaling factor of the fonts in the plots.[13]

```
R> library(testcorr)
R> cc.test(x, y, max.lag = 10, alpha = 0.05, lambda = 2.576, plot = TRUE,
+          table = TRUE, var.names = NULL, scale.font = 2)
```

We have the following testing outputs:





Tests for zero cross-correlation of x and y

Lag	CC	Stand. CB(95%)	Robust CB(95%)	Lag	t	p-value	t-tilde	p-value	Lag	HB	p-value	Q-tilde	p-value
-10	0.016	(-0.113, 0.113)	(-0.047, 0.047)	-10	0.281	0.779	0.677	0.498	-10	31.780	0.001	10.462	0.489
-9	0.013	(-0.113, 0.113)	(-0.058, 0.058)	-9	0.218	0.827	0.422	0.673	-9	31.698	0.000	10.003	0.440
-8	-0.007	(-0.113, 0.113)	(-0.127, 0.127)	-8	-0.122	0.903	-0.109	0.914	-8	31.649	0.000	9.825	0.365
-7	-0.081	(-0.113, 0.113)	(-0.157, 0.157)	-7	-1.407	0.159	-1.013	0.311	-7	31.634	0.000	9.813	0.278
-6	0.036	(-0.113, 0.113)	(-0.085, 0.085)	-6	0.630	0.529	0.839	0.401	-6	29.606	0.000	8.788	0.268
-5	0.107	(-0.113, 0.113)	(-0.141, 0.141)	-5	1.859	0.063	1.491	0.136	-5	29.201	0.000	8.083	0.232
-4	0.011	(-0.113, 0.113)	(-0.141, 0.141)	-4	0.195	0.845	0.157	0.876	-4	25.689	0.000	5.862	0.320
-3	-0.013	(-0.113, 0.113)	(-0.175, 0.175)	-3	-0.229	0.819	-0.147	0.883	-3	25.650	0.000	5.837	0.212
-2	0.157	(-0.113, 0.113)	(-0.197, 0.197)	-2	2.713	0.007	1.562	0.118	-2	25.597	0.000	5.815	0.121
-1	0.159	(-0.113, 0.113)	(-0.221, 0.221)	-1	2.746	0.006	1.405	0.160	-1	18.185	0.000	3.375	0.185
0	0.188	(-0.113, 0.113)	(-0.312, 0.312)	0	3.259	0.001	1.183	0.237	0	10.621	0.001	1.400	0.237
1	0.118	(-0.113, 0.113)	(-0.162, 0.162)	1	2.046	0.041	1.426	0.154	1	14.822	0.001	3.434	0.180
2	0.080	(-0.113, 0.113)	(-0.100, 0.100)	2	1.384	0.166	1.560	0.119	2	16.750	0.001	5.867	0.118
3	0.068	(-0.113, 0.113)	(-0.106, 0.106)	3	1.186	0.236	1.269	0.204	3	18.170	0.001	7.477	0.113
4	0.012	(-0.113, 0.113)	(-0.152, 0.152)	4	0.215	0.830	0.160	0.873	4	18.217	0.003	7.503	0.186
5	-0.069	(-0.113, 0.113)	(-0.158, 0.158)	5	-1.197	0.232	-0.857	0.391	5	19.673	0.003	8.238	0.221
6	0.067	(-0.113, 0.113)	(-0.125, 0.125)	6	1.167	0.243	1.056	0.291	6	21.062	0.004	9.353	0.228
7	0.099	(-0.113, 0.113)	(-0.213, 0.213)	7	1.718	0.086	0.914	0.361	7	24.084	0.002	10.188	0.252
8	-0.020	(-0.113, 0.113)	(-0.079, 0.079)	8	-0.343	0.732	-0.490	0.624	8	24.205	0.004	10.428	0.317
9	0.055	(-0.113, 0.113)	(-0.066, 0.066)	9	0.959	0.337	1.637	0.102	9	25.154	0.005	13.109	0.218
10	0.049	(-0.113, 0.113)	(-0.071, 0.071)	10	0.855	0.392	1.360	0.174	10	25.911	0.007	14.958	0.184

The left-hand side plot is graphing for maximum ± 10 lags, the sample cross-correlation $\widehat{\rho}_{xy,k}$ (`"CC"`), the standard and robust CB(95%). The right-hand side plot is graphing for maximum ± 10 lags, the cumulative test statistics $HB_{xy,m}$, $\widetilde{Q}_{xy,m}$ and their critical values at 5% significance level (`"cv(5%)"`). The table reports the results of the plots along with the p-values for all the statistics: standard $t_{xy,k}$ (`"t"`) and $HB_{xy,m}$ (`"HB"`) and robust $\widetilde{t}_{xy,k}$ (`"t-tilde"`) and $\widetilde{Q}_{xy,m}$ (`"Q-tilde"`). The columns of the table can each be extracted by adding `$lag`, `$cc`, `$scb`, `$rcb`, `$t`, `$pvt`, `$ttilde`, `$pvttilde`, `$hb`, `$pvhb`, `$qtilde`, `$pvqtilde` at the end of the function call.

From the left-hand side plot we can conclude that $H_0 : \rho_{xy,k} = 0$ is rejected at $\alpha = 5\%$ when $k = -2, -1, 0, 1$ and is not rejected at $\alpha = 5\%$ for $k \neq -2, -1, 0, 1$ using standard methods, but

[13] The default values are `alpha = 0.05`, `lambda = 2.576`, `plot = TRUE`, `table = TRUE`, `var.names = NULL` and `scale.font = 1`. Setting `scale.font = 2` is useful in order to upscale the fonts in the plots in order to export them as displayed here; the default value is suggested for viewing the plots in the Plots pane.

is not rejected at $\alpha = 5\%$ for any k using robust methods. From the right-hand side plot we can conclude that the cumulative hypothesis $H_0 : \rho_{xy,0} = ... = \rho_{xy,m} = 0$ is rejected at $\alpha = 5\%$ for all m using standard methods, but is not rejected at any m using robust methods. Subsequently, from the p-values in the table we find that using standard methods, $H_0 : \rho_{xy,k} = 0$ is rejected at $\alpha = 1\%$ when $k = -2, -1, 0$, at $\alpha = 5\%$ for $k = 1$, at $\alpha = 10\%$ when $k = -5, 7$ and is not rejected at $\alpha = 10\%$ for all $k \neq -5, -2, -1, 0, 1, 7$, whereas using robust methods it is not rejected at $\alpha = 10\%$ for any k. Using standard methods the cumulative hypothesis $H_0 : \rho_{xy,0} = ... = \rho_{xy,m} = 0$ is rejected at $\alpha = 0.1\%$ when $m = -10, ..., -1, 1, 2$ and at $\alpha = 1\%$ for $m = 0, 3, ..., 10$, whereas using robust methods it is not rejected at $\alpha = 10\%$ for any m. Overall, standard testing procedures show evidence of cross-correlation, although the series are uncorrelated from each other. The robust testing procedures provide the correct inference.

4. Testing zero Pearson correlation: `rcorr.test`

For multivariate series $\{x_{1t}, ..., x_{pt}\}$, given a sample $(x_{11}, ..., x_{1n}), ..., (x_{p1}, ..., x_{pn})$, the null hypothesis $H_0 : \rho_{x_i x_j} = 0$ of no correlation between variables $\{x_{it}, x_{jt}\}$ is tested at α significance level using the sample Pearson correlation $\widehat{\rho}_{x_i x_j}$ and the p-value of the robust t-type statistic $\widetilde{t}_{x_i x_j}$. This robust procedure is obtained from the $\widetilde{t}_{xy,k}$ test of Section 3 setting $x = x_i$, $y = x_j$ and $k = 0$.

Example

We provide an example to illustrate testing zero correlation between variables of a 4-dimensional series $\{x_{1t}, x_{2t}, x_{3t}, x_{4t}\}$ using the function `rcorr.test`. We use the simulated data from the series $\{x_t, y_t, z_t, u_t\}$ of Section 3. The pairs $\{x_t, u_t\}$ and $\{z_t, u_t\}$ are independent, $\{x_t, y_t\}$ and $\{y_t, z_t\}$ are uncorrelated but are dependent, while $\{x_t, z_t\}$ and $\{y_t, u_t\}$ are correlated. From the four series only $\{x_t\}$ and $\{u_t\}$ are i.i.d. We bind the series into a matrix.

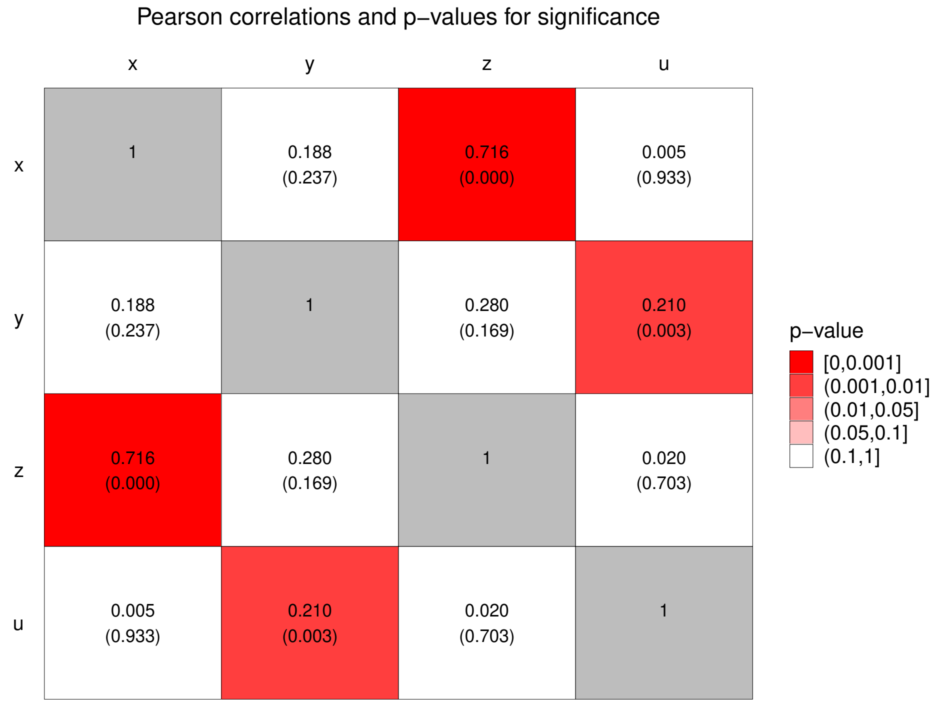
```
R> matx <- cbind(x, y, z, u)
```

We use the function `rcorr.test` to evaluate the results on testing. The plot is shown in the Plots pane and the tables are printed on the Console. We don't pass any variables' names and set to 1.5 the scaling factor of the fonts in the plot.[14]

```
R> library(testcorr)
R> rcorr.test(matx, plot = TRUE, table = TRUE, var.names = NULL,
+             scale.font = 1.5)
```

We have the following testing outputs:

[14]The default values are `plot = TRUE`, `table = TRUE`, `var.names = NULL` and `scale.font = 1`. Setting `scale.font = 1.5` is useful in order to upscale the fonts in the plot in order to export it as displayed here; the default value is suggested for viewing the plot in the Plots pane.




```
Matrix of Pearson correlations

      x     y     z     u
x     1 0.188 0.716 0.005
y 0.188     1 0.280 0.210
z 0.716 0.280     1 0.020
u 0.005 0.210 0.020     1
```

```
Matrix of p-values

      x     y     z     u
x       0.237 0.000 0.933
y 0.237       0.169 0.003
z 0.000 0.169       0.703
u 0.933 0.003 0.703
```

The plot is a heatmap of the sample Pearson correlations $\widehat{\rho}_{x_i x_j}$ among all pairs i, j of variables and their p-values (in parenthesis) for testing significance of correlation. Four shades of red, from dark to light, indicate significance at level $\alpha = 0.1\%, 1\%, 5\%, 10\%$, respectively, and white indicates non-significance at level $\alpha = 10\%$. The two tables report the results of the plot. The tables can each be extracted by adding `$pc`, `$pv` at the end of the function call.

From the p-values in the plot and the right-hand side table we can conclude that $H_0 : \rho_{xy} = 0$, $H_0 : \rho_{xu} = 0$, $H_0 : \rho_{yz} = 0$ and $H_0 : \rho_{zu} = 0$ are not rejected at $\alpha = 10\%$, $H_0 : \rho_{xz} = 0$ is rejected at $\alpha = 0.1\%$ and $H_0 : \rho_{yu} = 0$ is rejected at $\alpha = 1\%$. Overall, the robust testing procedure provides the correct inference. In contrast, the standard procedure[15] gives wrong inference when the series are uncorrelated but dependent. To demonstrate this, we use the function `rcorr` from the package **Hmisc** (Harrell Jr *et al.* 2019) to evaluate the sample Pearson correlations and their p-values for testing significance of correlation.

```
R> library(Hmisc)
R> print(format(round(rcorr(matx)$r, 3), nsmall = 3), quote = FALSE)
R> print(format(round(rcorr(matx)$P, 3), nsmall = 3), quote = FALSE)
```

[15]The standard procedure is implemented by `Hmisc::rcorr` and `stats::cor.test`. In these functions, the standard t-test differs slightly from that given in Section 3. In `Hmisc::rcorr` and `stats::cor.test` the statistic $t'_{x_i x_j} = \widehat{\rho}_{x_i x_j} \sqrt{(n-2)/(1-\widehat{\rho}^2_{x_i x_j})}$ and critical values from the t_{n-2} distribution are used, while in Section 3 we take $t_{x_i x_j} = \sqrt{n}\,\widehat{\rho}_{x_i x_j}$ and critical values from the $N(0,1)$ distribution. For big samples, they give very similar results under H_0. For example, in Section 3 we find p-value of 0.00112 in testing $H_0 : \rho_{xy} = 0$ with the standard t_{xy} test, while in the output from `Hmisc::rcorr` it is 0.00106 using the standard t'_{xy} test.

We have the following outputs:

```
      x     y     z     u
x 1.000 0.188 0.716 0.005
y 0.188 1.000 0.280 0.210
z 0.716 0.280 1.000 0.020
u 0.005 0.210 0.020 1.000
```

```
      x     y     z     u
x    NA 0.001 0.000 0.936
y 0.001    NA 0.000 0.000
z 0.000 0.000    NA 0.735
u 0.936 0.000 0.735    NA
```

From the p-values in the right-hand side table we can conclude that $H_0 : \rho_{xu} = 0$ and $H_0 : \rho_{zu} = 0$ are not rejected at $\alpha = 10\%$, $H_0 : \rho_{xz} = 0$, $H_0 : \rho_{yz} = 0$ and $H_0 : \rho_{yu} = 0$ are rejected at $\alpha = 0.1\%$ and $H_0 : \rho_{xy} = 0$ is rejected at $\alpha = 1\%$. Hence, using the standard procedure we wrongly conclude that the series $\{x_t\}$ with $\{y_t\}$ and $\{y_t\}$ with $\{z_t\}$ are correlated.

5. Testing i.i.d. property: `iid.test`

For a univariate series $\{x_t\}$, given a sample $x_1, ..., x_n$, the null hypothesis of the i.i.d. property is tested at lag $k = 1, 2, ...$ by verifying $H_0 : \rho_{x,k} = 0, \rho_{|x|,k} = 0$ or $H_0 : \rho_{x,k} = 0, \rho_{x^2,k} = 0$, using the $J_{x,|x|,k}$ and $J_{x,x^2,k}$ statistics.[16] The null hypothesis of the i.i.d. property at cumulative lags $m = 1, 2, ...$ is tested by verifying $H_0 : \rho_{x,k} = 0, \rho_{|x|,k} = 0, k = 1, ..., m$ or $H_0 : \rho_{x,k} = 0, \rho_{x^2,k} = 0, k = 1, ..., m$, using the $C_{x,|x|,m}$ and $C_{x,x^2,m}$ statistics. The following notation is used.

$$J_{x,|x|,k} = \frac{n^2}{n-k}(\widehat{\rho}^2_{x,k} + \widehat{\rho}^2_{|x|,k}), \quad C_{x,|x|,m} = \sum_{k=1}^{m} J_{x,|x|,k},$$
$$J_{x,x^2,k} = \frac{n^2}{n-k}(\widehat{\rho}^2_{x,k} + \widehat{\rho}^2_{x^2,k}), \quad C_{x,x^2,m} = \sum_{k=1}^{m} J_{x,x^2,k},$$

where $\widehat{\rho}_{x,k} = \widehat{\mathrm{corr}}(x_t, x_{t-k})$, $\widehat{\rho}_{|x|,k} = \widehat{\mathrm{corr}}(|x_t - \bar{x}|, |x_{t-k} - \bar{x}|)$, $\widehat{\rho}_{x^2,k} = \widehat{\mathrm{corr}}((x_t - \bar{x})^2, (x_{t-k} - \bar{x})^2)$ and $\bar{x} = n^{-1}\sum_{t=1}^{n} x_t$ with $\widehat{\mathrm{corr}}$ denoting the sample correlation estimate.

Applying the tests, at significance level α, $H_0 : \rho_{x,k} = 0, \rho_{|x|,k} = 0$ or $H_0 : \rho_{x,k} = 0, \rho_{x^2,k} = 0$ is rejected when $J_{x,|x|,k} > \chi^2_{2,\alpha}$ or $J_{x,x^2,k} > \chi^2_{2,\alpha}$. In turn, $H_0 : \rho_{x,k} = 0, \rho_{|x|,k} = 0, k = 1, ..., m$ or $H_0 : \rho_{x,k} = 0, \rho_{x^2,k} = 0, k = 1, ..., m$ is rejected when $C_{x,|x|,m} > \chi^2_{2m,\alpha}$ or $C_{x,x^2,m} > \chi^2_{2m,\alpha}$. Here, $\chi^2_{m,\alpha}$ stands for the upper α quantile of χ^2_m distribution.

Example

We provide an example to illustrate testing for the i.i.d. property of a univariate series $\{x_t\}$ using the function `iid.test`. We use the simulated data from the series $\{x_t\}$ of Section 3. The series $\{x_t\}$ is i.i.d.

We use the function `iid.test` to evaluate the results on testing for maximum 10 lags at significance level $\alpha = 5\%$. The plots are shown in the Plots pane and the table is printed on the Console. We don't pass any variable's name and set to 2 the scaling factor of the fonts in the plots.[17,18]

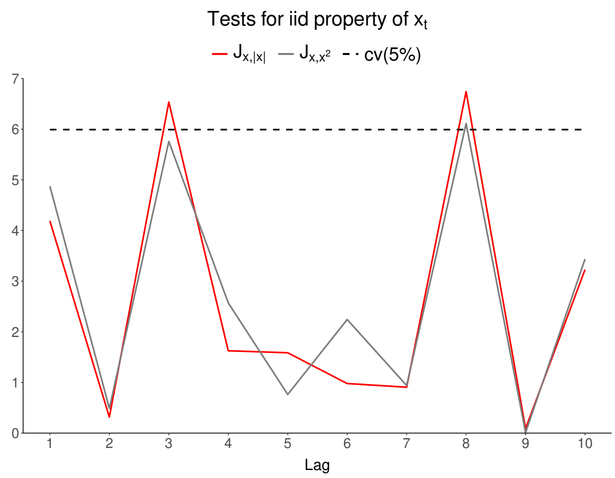
[16] Notation: $\rho_{x,k} = \mathrm{corr}(x_t, x_{t-k})$, $\rho_{|x|,k} = \mathrm{corr}(|x_t - \mu|, |x_{t-k} - \mu|)$, $\rho_{x^2,k} = \mathrm{corr}((x_t - \mu)^2, (x_{t-k} - \mu)^2)$ and $\mu = Ex_t$.

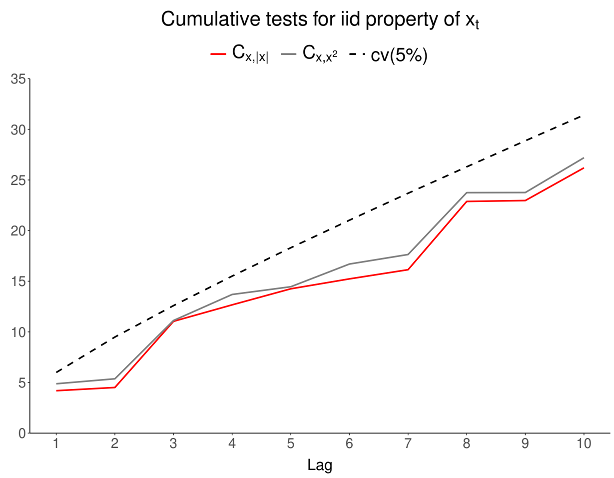
[17] The first letter of the variable's name is used as subscript instead of x in the statistics when `var.name` is not `NULL`.

[18] The default values are `alpha = 0.05`, `plot = TRUE`, `table = TRUE`, `var.name = NULL` and `scale.font = 1`. Set-

```
R> library(testcorr)
R> iid.test(x, max.lag = 10, alpha = 0.05, plot = TRUE, table = TRUE,
+           var.name = NULL, scale.font = 2)
```

We have the following testing outputs:





Tests for i.i.d. property of x

Lag	J[x,|x|]	p-value	J[x,x²]	p-value	Lag	C[x,|x|]	p-value	C[x,x²]	p-value
1	4.189	0.123	4.876	0.087	1	4.189	0.123	4.876	0.087
2	0.317	0.853	0.489	0.783	2	4.507	0.342	5.365	0.252
3	6.534	0.038	5.757	0.056	3	11.041	0.087	11.122	0.085
4	1.626	0.444	2.571	0.277	4	12.666	0.124	13.692	0.090
5	1.586	0.452	0.763	0.683	5	14.252	0.162	14.455	0.153
6	0.979	0.613	2.243	0.326	6	15.231	0.229	16.698	0.161
7	0.906	0.636	0.940	0.625	7	16.138	0.305	17.638	0.224
8	6.741	0.034	6.110	0.047	8	22.878	0.117	23.748	0.095
9	0.090	0.956	0.012	0.994	9	22.968	0.192	23.759	0.163
10	3.228	0.199	3.436	0.179	10	26.196	0.159	27.195	0.130

The plots are graphing for maximum 10 lags, the test statistics $J_{x,|x|,k}$, $J_{x,x^2,k}$ (left), the cumulative test statistics $C_{x,|x|,m}$, $C_{x,x^2,m}$ (right) and their critical values at 5% significance level (`"cv(5%)"`). The table reports the results of the plots along with the p-values for all the statistics: $J_{x,|x|,k}$ (`"J[x,|x|]"`), $J_{x,x^2,k}$ (`"J[x,x`^{2}`]"`), $C_{x,|x|,m}$ (`"C[x,|x|]"`) and $C_{x,x^2,m}$ (`"C[x,x`^{2}`]"`). The columns of the table can each be extracted by adding `$lag`, `$jab`, `$pvjab`, `$jsq`, `$pvjsq`, `$cab`, `$pvcab`, `$csq`, `$pvcsq` at the end of the function call.

From the left-hand side plot we can conclude that $H_0 : \rho_{x,k} = 0, \rho_{|x|,k} = 0$ is not rejected at $\alpha = 5\%$ for any k except $k = 3, 8$ or $H_0 : \rho_{x,k} = 0, \rho_{x^2,k} = 0$ is not rejected at $\alpha = 5\%$ for any k except $k = 8$. From the right-hand side plot we can conclude that the cumulative hypothesis $H_0 : \rho_{x,k} = 0, \rho_{|x|,k} = 0, k = 1, ..., m$ or $H_0 : \rho_{x,k} = 0, \rho_{x^2,k} = 0, k = 1, ..., m$ is not rejected at $\alpha = 5\%$ for any m. Subsequently, from the p-values in the table we find that $H_0 : \rho_{x,k} = 0, \rho_{|x|,k} = 0$ is rejected at $\alpha = 5\%$ for $k = 3, 8$ and is not reject at $\alpha = 10\%$ when $k \neq 3, 8$ or $H_0 : \rho_{x,k} = 0, \rho_{x^2,k} = 0$ is rejected at $\alpha = 5\%$ for $k = 8$ and at $\alpha = 10\%$ for $k = 1, 3$ and is not rejected at $\alpha = 10\%$ when $k \neq 1, 3, 8$. The cumulative hypothesis $H_0 : \rho_{x,k} = 0, \rho_{|x|,k} = 0, k = 1, ..., m$ is rejected at $\alpha = 10\%$ for $m = 3$ and is not rejected at $\alpha = 10\%$ when $m \neq 3$ or $H_0 : \rho_{x,k} = 0, \rho_{x^2,k} = 0, k = 1, ..., m$ is

ting `scale.font = 2` is useful in order to upscale the fonts in the plots in order to export them as displayed here; the default value is suggested for viewing the plots in the Plots pane.

rejected at $\alpha = 10\%$ for $m = 1, 3, 4, 8$ and is not rejected at $\alpha = 10\%$ for $m \neq 1, 3, 4, 8$. Overall, the testing procedures provide the correct inference.

6. Remarks

The theory and Monte Carlo study in Dalla *et al.* (2020) suggest that:

(i) In testing for autocorrelation the series needs to have constant mean.

(ii) In testing for cross-correlation each of the series needs to have constant mean and to be serially uncorrelated when applying the portmanteau type statistics or at least one when applying the t-type tests.

(iii) In testing for Pearson correlation at least one of the series needs to have constant mean and to be serially uncorrelated.

(iv) For relatively large lag it may happen that the robust portmanteau statistic is negative. In such a case, missing values (`NA`) are reported for the statistic and its p-value.

(v) The values $\lambda = 1.96, 2.576$ are good candidates for the threshold in the robust portmanteau statistics, with $\lambda = 2.576$ performing better at relatively large lags.

Acknowledgments

Dalla acknowledges financial support from ELKE-EKPA. Phillips acknowledges support from the Kelly Fund at the University of Auckland and a KLC Fellowship at Singapore Management University.

References

Boshnakov GN, Halliday J (2019). ***sarima**: Simulation and Prediction with Seasonal ARIMA Models*. R package version 0.8.1, URL `https://CRAN.R-project.org/package=sarima`.

Dalla V, Giraitis L, Phillips PCB (2020). "Robust Tests for White Noise and Cross-Correlation." Cowles Foundation, Discussion Paper No. 2194, URL `https://cowles.yale.edu/sites/default/files/files/pub/d21/d2194-r.pdf`.

Harrell Jr FE, with contributions from Charles Dupont, *et al.* (2019). ***Hmisc**: Harrell Miscellaneous*. R package version 4.2-0, URL `https://CRAN.R-project.org/package=Hmisc`.

Hart A, Martínez S (2018). ***spgs**: Statistical Patterns in Genomic Sequences*. R package version 1.0-2, URL `https://CRAN.R-project.org/package=spgs`.

Haugh LD, Box GEP (1977). "Identification of Dynamic Regression (Distributed Lag) Models Connecting Two Time Series." *Journal of the American Statistical Association*, **72**(357), 121–130. `doi:10.1080/01621459.1977.10479920`.

Ljung GM, Box GEP (1978). "On a Measure of Lack of Fit in Time Series Models." *Biometrika*, **65**(2), 297–303. doi:10.1093/biomet/65.2.297.

Mahdi E, McLeod AI (2018). ***portes**: Portmanteau Tests for Univariate and Multivariate Time Series Models*. R package version 3.0, URL https://CRAN.R-project.org/package=portes.

Pearson K (1896). "Mathematical Contributions to the Theory of Evolution. III. Regression, Heredity, and Panmixia." *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, **187**, 253–318. doi:10.1098/rsta.1896.0007.

R Core Team (2019). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Wickham P (2012). ***normwhn.test**: Normality and White Noise Testing*. R package version 1.0, URL https://CRAN.R-project.org/package=normwhn.test.

Yule UG (1926). "Why Do we Sometimes Get Nonsense-Correlations between Time-Series? A Study in Sampling and the Nature of Time-Series." *Journal of the Royal Statistical Society*, **89**(1), 1–63. doi:10.2307/2341482.

Affiliation:

Violetta Dalla
Department of Economics
National and Kapodistrian University of Athens
Athens 10559, Greece
E-mail: vidalla@econ.uoa.gr
URL: http://en.econ.uoa.gr/contact-info-dalla-violeta.html

Liudas Giraitis
School of Economics and Finance
Queen Mary University of London
London E1 4NS, UK
E-mail: l.giraitis@qmul.ac.uk
URL: https://www.qmul.ac.uk/sef/staff/liudasgiraitis.html

Peter C. B. Phillips
Department of Economics
Yale University
New Haven CT 06520-8281, USA
E-mail: peter.phillips@yale.edu
URL: https://economics.yale.edu/people/faculty/peter-phillips