Psicológica (2010), 31, 357-381.

Lag-one autocorrelation in short series: Estimation and hypotheses testing

Antonio Solanas[*1,3], Rumen Manolov[1], and Vicenta Sierra[2]

[1] *University of Barcelona, Spain;*
[2] *ESADE-Ramon Llull University, Barcelona, Spain;*
[3] *Institute for Research in Brain, Cognition, and Behavior (IR3C)*

In the first part of the study, nine estimators of the first-order autoregressive parameter are reviewed and a new estimator is proposed. The relationships and discrepancies between the estimators are discussed in order to achieve a clear differentiation. In the second part of the study, the precision in the estimation of autocorrelation is studied. The performance of the ten lag-one autocorrelation estimators is compared in terms of Mean Square Error (combining bias and variance) using data series generated by Monte Carlo simulation. The results show that there is not a single optimal estimator for all conditions, suggesting that the estimator ought to be chosen according to sample size and to the information available on the possible direction of the serial dependence. Additionally, the probability of labelling an actually existing autocorrelation as statistically significant is explored using Monte Carlo sampling. The power estimates obtained are quite similar among the tests associated with the different estimators. These estimates evidence the small probability of detecting autocorrelation in series with less than 20 measurement times.

The present study focuses on autocorrelation estimators reviewing most of them and proposing a new one. Hypothesis testing is also explored and discussed as the statistical significance of the estimates may be of interest. These topics are relevant for methodological and behavioural sciences, since they have impact on the techniques used for assessing intervention effectiveness.

---

[*] This research was supported by the *Comissionat per a Universitats i Recerca del Departament d'Innovació, Universitats i Empresa* of the *Generalitat de Catalunya* and the European Social Fund. Correspondence concerning this article should be addressed to Antonio Solanas, Departament de Metodologia de les Ciències del Comportament, Facultat de Psicologia, Universitat de Barcelona, Passeig de la Vall d'Hebron, 171, 08035-Barcelona, Spain. Phone number: +34933125076. Electronic mail may be sent to Antonio Solanas at antonio.solanas@ub.edu

It has to be taken into consideration that the previous decades' controversy on the existence of autocorrelation in behavioural data (Busk & Marascuilo, 1988; Huitema, 1985; 1988; Sharpley & Alavosius, 1988; Suen & Ary, 1987) was strongly related to the properties of the autocorrelation estimators. The evidence on the presence of serial dependence (Matyas & Greenwood, 1997; Parker, 2006) has led to exploring the effects of violating the assumptions of independence of several widely used procedures. In this relation, liberal Type I error rates have been obtained in presence of positive serial dependence for traditional analysis of variance (Scheffé, 1959) and its modifications (Toothaker, Banz, Noble, Camp, & Davis, 1983). Additionally, randomization tests – a procedure that does not explicitly assume independence (Edgington, & Onghena, 2007) – have shown to be affected by positive autocorrelation both in terms in reducing statistical power (Ferron & Ware, 1995) and, more recently, in distorting Type I error rates (Manolov & Solanas, 2009). The independence of residuals required by regression analysis (Weisberg, 1980) has resulted in proposing that after fitting the regression model, a statistically significant autocorrelation in the errors has to be eliminated prior to interpreting the regression coefficients. For instance, generalized least squares procedures such as the one proposed by Simonton (1977) and the Cochrane-Orcutt and Prais-Winsten versions require estimating the autocorrelation of the residuals. Imprecisely estimated serial dependence may lead to elevated Type I error rates when assessing intervention effectiveness in short series.

Autoregressive integrated moving average (ARIMA) modeling has also been proposed for dealing with sequentially related data (Box & Jenkins, 1970). This procedure includes an initial step of model identification including autocorrelation estimation prior to controlling it and determining the efficacy of the interventions. However, it has been shown that serial dependence distorts the performance of ARIMA in short series (Greenwood & Matyas, 1990). Unfortunately, the required amount of measurements is not frequent in applied psychological studies and, moreover, it does not ensure correct model identification (Velicer & Harrop, 1983).

Several investigations (Arnau & Bono, 2001; DeCarlo & Tryon, 1993; Huitema & McKean, 1991, 2007a, b; Matyas & Greenwood, 1991; McKean & Huitema, 1993) have carried out Monte Carlo simulation comparisons of autocorrelation estimators for different lags. These studies have shown that estimation and hypothesis testing are both problematic in short data series. Most of the estimators studied had considerable bias and were scarcely efficient for short series. As regards the asymptotic test based on Bartlett's (1946) proposal, it proved to be unacceptable. These topics have to be taken into consideration when using widespread statistical packages, as they incorporate asymptotic results in their algorithms, making the correspondence between empirical and nominal Type I error rates dubious and compromising statistical power. Therefore, basic and applied researchers should know which estimators are incorporated in the statistical software, their mathematical expression and the asymptotic approximation used for testing hypotheses.

The main objectives of the present study were: a) describe several lag-one autocorrelation estimators, presenting the expressions for their calculus; b) propose a new estimator and test it in comparison with the previously developed estimators in terms of bias and Mean Square Error (hereinafter, MSE); c) estimate the statistical power of the tests associated with the ten estimators and based on Monte Carlo sampling.

Lag-one autocorrelation estimators

The rationale behind the present review can be found in the lack of an integrative compilation of autocorrelation estimators. Their correct identification is necessary in order to avoid confusions – for instance, Cox's (1966) research seemed to centre on the *conventional* estimator, while in fact it was the *modified* one (Moran, 1970), both being presented subsequently.

Conventional estimator

Although there is a great diversity of autoregressive parameter estimators, the most frequently utilised one in social and behavioural sciences is the *conventional* one (as referred to by Huitema & McKean, 1991). This estimator is defined by the following expression:

$$r_1 = \frac{\sum_{i=1}^{n-1} (x_i - \overline{x})(x_{i+1} - \overline{x})}{\sum_{i=1}^{n} (x_i - \overline{x})^2}$$

Its mathematical expectancy, presented in Kendall and Ord (1990), shows that its bias approximates $-(1 + 4\rho)/n$ for long series, where ρ is the autoregressive parameter and n is the series length. It has been demonstrated (Moran, 1948) that in independent processes $-n^{-1}$ is an exact result for r_1's bias without assuming the normality of the random term. As regards the variance of r_1, Bartlett's (1946) equation is commonly used, although several investigations (Huitema & McKean, 1991; Matyas & Greenwood, 1991) have shown that it does not approximate sufficiently the data obtained through Monte Carlo simulation. The lack of matching between nominal and empirical Type I error rates and the inadequate power of the asymptotic statistical test reported by previous studies may be due to the bias of the estimator and the asymmetry of the sampling distribution.

Modified estimator

Orcutt (1948) proposed the following estimator of autoregressive parameters:

$$r_1^* = \frac{n}{n-1} \frac{\sum_{i=1}^{n-1} (x_i - \bar{x})(x_{i+1} - \bar{x})}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$$

Hereinafter, this estimator will be referred to as the *modified* estimator as it consists in a linear modification of the conventional estimator presented above. On the basis of its mathematical expectancy described by Marriott and Pope (1954) it can be seen that the bias of the *modified* estimator approximates $-(1 + 3\rho)/n$ for long series and, thus, it is not identical to the one of the *conventional* estimator, as it has been assumed (Huitema & McKean, 1991). The differences in independent processes bias reported by Moran (1948) and Marriott and Pope (1954) can be due to the asymmetry of the sampling distribution of the estimator. This puts in doubt the utility of the mathematical expectancy as a bias criterion (Kendall, 1954). Moran (1967) demonstrated that $Var(r_1^*)$ depends on the shape of the distribution of the random term.

Cyclic estimator

A *cyclic* estimator for different lag autocorrelations was investigated by Anderson (1942), although it was previously proposed by H. Hotelling (Moran, 1948). It is defined as:

$$r_1^c = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(x_{i+1} - \bar{x})}{\sum_{i=1}^{n} (x_i - \bar{x})^2}, x_{n+1} = x_1$$

In independent processes, Anderson (1942) derived an exact distribution of the lag-one estimator for several series lengths. The distribution is highly asymmetric in short series and, according to Kendall (1954), in those cases bias should not be determined by means of procedures based on the mathematical expectancy.

Exact estimator

The expression for the *exact* estimator (Kendall, 1954) corresponds to the one generally used for calculating the correlation coefficient:

$$r_1^e = \frac{A}{\sqrt{BC}}, \text{ where}$$

$$A = \frac{1}{n-1}\sum_{i=1}^{n-1} x_i x_{i+1} - \frac{1}{(n-1)^2}\left(\sum_{i=1}^{n-1} x_i\right)\left(\sum_{i=1}^{n-1} x_{i+1}\right)$$

$$B = \frac{1}{n-1}\sum_{i=1}^{n-1} x_i^2 - \frac{1}{(n-1)^2}\left(\sum_{i=1}^{n-1} x_i\right)^2$$

$$C = \frac{1}{n-1}\sum_{i=1}^{n-1} x_{i+1}^2 - \frac{1}{(n-1)^2}\left(\sum_{i=1}^{n-1} x_{i+1}\right)^2$$

Mathematical-expectancy-based procedures led Kendall (1954) to the attainment of the bias of the estimator in independent processes: approximately $-1/(n-1)$ for long series.

C statistic

The *C* statistic was developed by Young (1941) in order to determine if data series are random or not. Although it has been commented and tested for assessing intervention effectiveness (Crosbie, 1989; Tryon, 1982; 1984), DeCarlo and Tryon (1993) demonstrated that the *C* statistic is an estimator of lag-one autocorrelation, despite the fact it does not perform as expected in short data series. The *C* statistic can be obtained through the following expression:

$$C = 1 - \frac{\sum_{i=1}^{n-1}(x_i - x_{i+1})^2}{2\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

Fuller's estimator

Fuller (1976) proposed an estimator supposed to correct the *conventional* estimator's bias, especially for short series. The following expression represents what we refer to as the *Fuller* estimator:

$$r_1^f = r_1 + \frac{1}{(n-1)}\left(1 - r_1^2\right)$$

Least squares estimators

Tuan (1992) presents two least squares estimators, whose lag-one formulae can be expressed in the following manner:

Least squares estimator:

$$r_1^{ls} = \frac{\sum_{i=1}^{n-1}(x_i - \bar{x})(x_{i+1} - \bar{x})}{\sum_{i=1}^{n-1}(x_i - \bar{x})^2}$$

Least squares *forward-backward* estimator:

$$r_1^{fb} = \frac{\sum_{i=1}^{n-1}(x_i - \bar{x})(x_{i+1} - \bar{x})}{\frac{1}{2}(x_1 - \bar{x})^2 + \sum_{i=2}^{n-1}(x_i - \bar{x})^2 + \frac{1}{2}(x_n - \bar{x})^2}$$

In the first expression, in the denominator there are only $n-1$ terms, as the information about the last data point is omitted. The second expression has n terms in its denominator, where the additional term arises from an averaged deviate of the initial and final data points.

Translated estimator

The r_1^+ estimator was proposed by Huitema and McKean (1991):

$$r_1^+ = r_1 + \frac{1}{n}$$

Throughout this article it will be referred to as the *translated* estimator, as it performs a translation over the *conventional* estimator in order to correct part of the n^{-1} bias. It can be demonstrated that $Bias(r_1^+)$ is approximately $-(4\rho)/n$.

Other autocorrelation estimators

It is practically impossible for a single investigation to assess all existing methods to estimate autocorrelation. The present study includes only the estimators which are common in behavioural sciences literature and in statistical packages, omitting, for instance, estimator r_1' fitted by the bias (Arnau, 1999; Arnau & Bono, 2001). Additionally, the estimators proposed by Huitema and McKean (1994c) and the *jackknife* estimator (Quenouille, 1949) were not included in this study, since they are not very efficient despite the bias reduction they perform. In fact, both the jackknife and the bootstrap methods are not estimators themselves but can rather be applied to any estimator in order to reduce its bias, as has already been done (Huitema & McKean, 1994a; McKnight, McKean, & Huitema, 2000).

The *maximum likelihood* estimator is obtained resolving a cubic equation and assuming an independent and normal distribution of the errors. There is an expression of this estimator (Kendall & Ord, 1990) which would be more easily incorporated in statistical software, but it has not been contrasted in any other article, nor do the authors justify the simplification they propose.

The δ-recursive estimator

The present investigation proposes a new lag-one autocorrelation estimator, referred to as the *δ-recursive* estimator, which is defined as follows:

$$r_1^{\delta} = \left(r_1 + \frac{1}{n}\right)\left(1 + \frac{\delta}{n}\right) = r_1^{+}\left(1 + \frac{\delta}{n}\right), \quad \delta \geq 0 .$$

In the expression above, r_1 is the *conventional* estimator, r_1^+ is the *translated* estimator, n corresponds to the length of the data series, and δ is a constant for bias correction. This expression illustrates the close relationship between the *translated* and the proposed estimator, highlighting their equivalence when δ is equal to zero. As it can be seen, an additional correction is introduced to the *translated* estimator, since it is only unbiased for independent data series. Therefore, the objective of the *δ-recursive* estimator is to maintain the desirable properties of r_1^+ for $\rho_1 = 0$ and to reduce bias for $\rho_1 \neq 0$. This reduction of bias is achieved by means of the acceleration constant δ; a greater value of δ implies a greater reduction in bias, always keeping in mind that bias is also reduced when more measurements (n) are available. However, it has to be taken into account that $\mathrm{Var}(r_1^{\delta}) = \mathrm{Var}[r_1^{+}(1+ \delta/n)] = (1+ \delta/n)^2\, \mathrm{Var}(r_1^{+})$ and, thus, for greater values of the constant, the proposed estimator becomes less efficient than the translated one. Therefore, the value of *δ* has to be chosen in a way to reduce the MSE and not only bias, in order the proposed estimator to be useful.

Some analytical and asymptotical results have been derived for the *δ-recursive* first order estimator:

$$E(r_1^{\delta}|\rho_1) = (n+\delta)\left(\frac{E(r_1|\rho_1)}{n} + \frac{1}{n^2}\right)$$

$$Var(r_1^{\delta}|\rho_1) = \left(\frac{(n+\delta)}{n}\right)^2 Var(r_1|\rho_1)$$

$$Bias(r_1^{\delta}|\rho_1) = (n+\delta)\left(\frac{E(r_1|\rho_1)}{n} + \frac{1}{n^2}\right) - \rho_1$$

Regarding the asymptotic distribution of the *δ-recursive* estimator in independent processes,

$$r_1^{\delta} \to \mathbf{N}\left(0; \frac{(n-2)^2(n+\delta)^2}{n^4(n-1)}\right).$$

Although there is a considerable matching between the theoretical and empirical sampling distributions for 50 data points, preliminary studies suggest that 100 measurement points are necessary.

Monte Carlo simulation: Mean Square Error

Method

The first experimental section of the current investigation consists in a comparison between the different lag-one autocorrelation estimators in terms of a precision indicator like MSE, which contains information about both bias and variance. This measure was chosen as it has been suggested to be appropriate for describing both biased and unbiased estimators (Spanos, 1987) and for comparing between estimators (Jenkins & Watts, 1968).

The computer-intensive technique utilised was Monte Carlo simulation, which is the optimal choice when the population distribution (i.e., the value of the autoregressive parameter and random variable distribution) is known (Noreen, 1989). Data series with ten different lengths (*n* = 5, 6, 7, 8, 9, 10, 15, 20, 50, and 100) were generated using a first order autoregressive model of the form $\mathbf{e_t} = \boldsymbol{\rho_1}\ \mathbf{e_{t-1}} + \mathbf{u_t}$ testing nineteen levels of the lag-one autocorrelation (ρ_1): −.9(.1).9. This model and these levels of serial dependence are the most common one in studies on autocorrelation estimation (e.g., Huitema & McKean, 1991, 1994b; Matyas & Greenwood, 1991). The error term followed three different distribution shapes with the same mean (zero) and the same standard deviation (one). Nonnormal distributions were included apart from the typically used normal distribution, due to the evidence that normal distributions may not represent sufficiently well behavioural data in some cases (Bradley, 1977; Micceri, 1989). Nonnormal distributions have already been studied in other contexts (Sawilowsky & Blair, 1992). In the present research we chose a uniform distribution in order to study the importance of kurtosis (a rectangular distribution is more platykurtic than the normal one with a γ_2 value of −1.2), specifying the α and β (i.e., minimum and maximum) parameters to be equal to −1.7320508075688773 and 1.7320508075688773, respectively, in order to obtain the abovementioned mean and variance. A negative exponential distribution was employed to explore the effect of skewness, as this type of distribution is asymmetrical in contrast to the Gaussian distribution, with a γ_1 value of 2. Zero mean and unity standard deviation were achieved simulating a one-parameter distribution ($\theta = 0$) with scale parameter σ equal to 1 and subtracting one from the data.

For each of the 570 experimental conditions 300,000 samples were generated using Fortran 90 and the NAG libraries *nag_rand_neg_exp*, *nag_rand_normal*, and *nag_rand_uniform*. We verified the correct simulation process comparing the theoretical results available in the scientific literature with the estimators' mean and variance computed from simulated data.

Prior to comparing the ten estimators, we carried out a preliminary study on the optimal value of δ for different series lengths in terms of minimizing MSE across all levels of autocorrelation from −.9 to .9. Monte Carlo simulations involving 300,000 iterations per experimental condition suggest that the optimal δ depends on the errors' distribution shape. Nevertheless, as applied researchers are not likely to know the errors' distribution, we chose a δ that is suitable for the three distributional shapes studied. For series lengths from 5 to 9 the optimal value resulted to be 0 and, thus, the MSE values for the δ-recursive estimator are the same as for the translated estimator. For $n = 10$ the δ constant was set to .4, for $n = 15$ to .9, and for $n = 20$ to 1.2. For longer series, lower MSE values were obtained for δ ranging from .7 to 1.5. As there was practically no difference between those values for series with 50 and 100 data points, δ was set to 1 – the only integer in that interval.

Results

The focus of this section is on intermediate levels of autocorrelation (between −.6 and .6) as those have been found to be more frequent in single-case data (Matyas & Greenwood, 1997; Parker, 2006). On the other hand, the results for shorter data series will be emphasised, as those appear to be more common in behavioural data (Huitema, 1985).

There is an exponential decay of MSE with the increase of the series length and the differences between the estimators are also reduced to minimum for *n* > *20*, as Figure 1 shows.

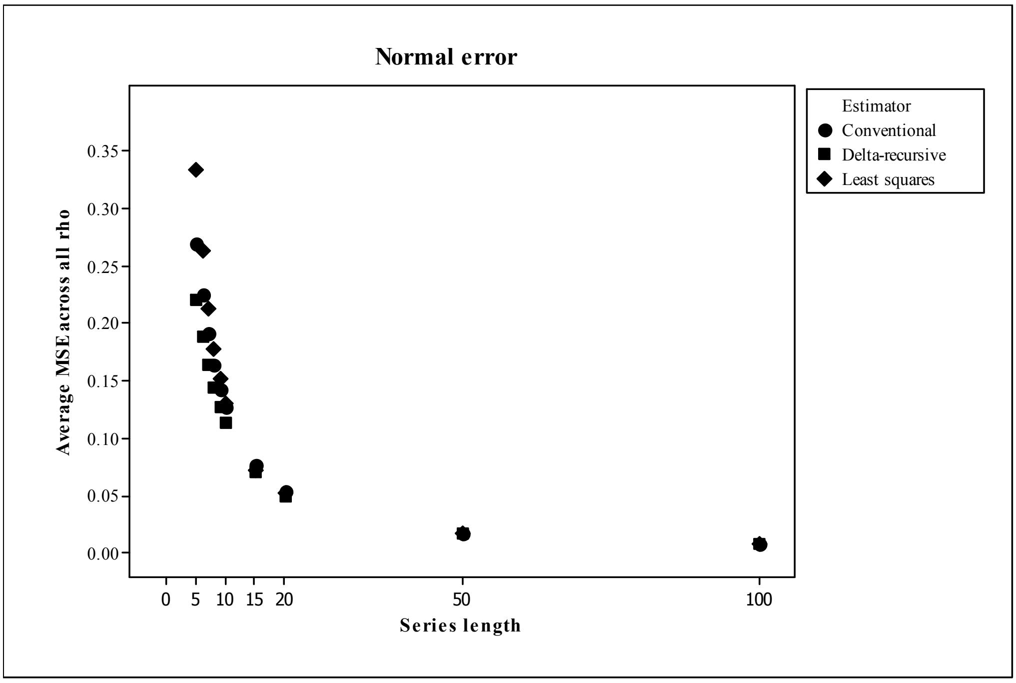

Figure 1. Example of the decrease of MSE (averaged across all ρ_1) for three autocorrelation estimators in series with normally distributed error.

The average MSE over all values of ρ_1 studied can be taken as a general indicator of the performance of the estimators. This information can also be useful for an applied researcher who has to choose an autocorrelation estimator and has no clue on the possible direction and level of serial dependence. The *translated* estimator shows lower MSE for series of length 5 to 9, while for $n \geq 10$ it is better to use the *Fuller,* the *translated*, or the *δ-recursive* estimators, which show practically equivalent MSE values, outperforming the remaining estimators (see Table 1). The *δ-recursive* estimator performed slightly better than any of the estimators tested for $n \geq 15$ series. It is important to remark that the *conventional* estimator, commonly used in the behavioural sciences, is not the most adequate one in terms of MSE.

It has to be highlighted that there is a notable divergence between the best performers for negative and positive serial dependence. As regards $\rho_1 = -.3$ (see Table 1), the *conventional* and the *cyclic* estimators show a better performance for $n \leq 20$. For $\rho_1 = .0$ (see Table 2), the estimators with lower MSE are the *translated*, *Fuller*, and the *δ-recursive*. For positive values of the autoregressive parameter (Table 2), the same three estimators and the *C* statistic excel.

When focusing on bias, as one of the components of MSE, the *conventional* and the *cyclic* estimators prove to be less biased for low negative autocorrelation (Table 3), while the *translated*, the *C* statistic, and the *δ-recursive* estimators are unbiased for independent data series (Table 4). Table 4 also contains the information about some positive values of the autoregressive parameter. For $\rho_1 = .3$, the bias of the *Fuller*, the *translated*, the *C* statistic, and the *δ-recursive* estimators is half the bias of the remaining estimators for $5 \leq n \leq 10$. For higher positive serial dependence, the aforementioned four estimators are once again the less biased ones. The proposed *δ-recursive* estimator is the less biased one for positive autocorrelation and series with 10 and 15 data points, cases in which δ was set to .4 and .9, respectively.

As regards the relevance of errors' distribution, Figure 2 illustrates the general finding that MSE tends to be somewhat smaller when the errors follow a negative (i.e., positive asymmetric) exponential distribution and greater when they are uniformly distributed.

Table 1. Mean square error of the ten lag-one autocorrelation estimators in series with different lengths. Average: bias averaged across −.9 ≤ ρ_1 ≤.9.

Estimators	Auto-correlation	SERIES LENGTH								
		5	10	15	5	10	15	5	10	15
		Exponential errors			Normal errors			Uniform errors		
Conven-tional	Average	.257	.116	.071	.270	.127	.077	.281	.133	.080
	−.6.	.102	.067	.047	.107	.069	.047	.113	.073	.049
	−.3.	.081	.061	.047	.086	.066	.049	.092	.071	.052
Modified	Average	.297	.119	.070	.315	.131	.077	.331	.137	.080
	−.6.	.122	.072	.049	.127	.073	.049	.134	.077	.051
	−.3.	.135	.077	.054	.143	.082	.057	.152	.088	.060
Cyclic	Average	.308	.127	.074	.322	.138	.071	.334	.145	.085
	−.6.	.092	.069	.048	.103	.072	.049	.111	.077	.051
	−.3.	.094	.066	.049	.103	.073	.052	.111	.078	.056
Exact	Average	.335	.119	.069	.345	.129	.075	.356	.135	.079
	−.6.	.151	.070	.047	.146	.069	.046	.154	.074	.049
	−.3.	.188	.080	.054	.182	.083	.056	.189	.088	.060
C statistic	Average	.232	.112	.069	.239	.117	.076	.250	.122	.075
	−.6.	.238	.116	.072	.237	.112	.068	.245	.114	.069
	−.3.	.153	.089	.062	.154	.089	.061	.158	.092	.062
Fuller	Average	.211	.103	.065	.225	.113	.070	.237	.118	.074
	−.6.	.221	.100	.062	.231	.103	.063	.239	.107	.065
	−.3.	.138	.077	.054	.147	.084	.058	.154	.089	.061
Least Squares	Average	.318	.122	.070	.334	.131	.072	.345	.137	.079
	−.6.	.130	.073	.048	.139	.074	.048	.143	.078	.050
	−.3.	.149	.081	.055	.159	.084	.057	.165	.089	.060
Forward-Backward	Average	.288	.117	.068	.306	.128	.075	.320	.135	.079
	−.6.	.109	.068	.046	.114	.068	.046	.121	.073	.049
	−.3.	.123	.074	.052	.130	.079	.055	.139	.085	.059
Translated	Average	.209	.103	.065	.221	.113	.081	.231	.119	.074
	−.6.	.205	.098	.062	.214	.101	.062	.221	.104	.063
	−.3.	.114	.072	.051	.121	.077	.055	.127	.082	.058
δ-recursive	Average	.209	.103	.064	.221	.113	.071	.231	.119	.073
	−.6.	.205	.097	.060	.214	.099	.060	.221	.105	.065
	−.3.	.114	.075	.055	.121	.081	.059	.127	.087	.063

Table 2. Mean square error of the ten different lag-one autocorrelation estimators in series with different lengths.

Estimators	Auto-correlation	SERIES LENGTH								
		5	10	15	5	10	15	5	10	15
		Exponential errors			Normal errors			Uniform errors		
Conven-tional	0	.123	.072	.052	.130	.081	.058	.137	.087	.062
	.3	.242	.106	.066	.253	.119	.076	.264	.125	.079
	.6	.464	.177	.097	.485	.195	.108	.506	.203	.112
Modified	0	.192	.089	.059	.203	.100	.067	.214	.107	.071
	.3	.310	.116	.069	.325	.132	.081	.342	.140	.085
	.6	.517	.173	.092	.545	.194	.104	.573	.202	.108
Cyclic	0	.157	.081	.055	.167	.091	.063	.176	.097	.066
	.3	.301	.117	.070	.315	.132	.081	.329	.140	.085
	.6	.571	.192	.102	.588	.212	.114	.609	.220	.118
Exact	0	.255	.095	.061	.258	.104	.068	.267	.110	.071
	.3	.371	.122	.071	.381	.136	.081	.393	.143	.085
	.6	.556	.174	.091	.578	.193	.103	.595	.201	.107
C statistic	0	.122	.077	.055	.125	.081	.058	.128	.083	.060
	.3	.156	.083	.055	.160	.091	.062	.167	.095	.064
	.6	.269	.118	.070	.283	.132	.079	.303	.139	.082
Fuller	0	.100	.064	.048	.109	.074	.055	.116	.080	.058
	.3	.124	.071	.049	.135	.084	.059	.146	.090	.062
	.6	.242	.115	.070	.259	.131	.080	.278	.138	.083
Least Squares	0	.215	.096	.062	.226	.104	.068	.234	.109	.071
	.3	.344	.126	.072	.355	.137	.082	.369	.143	.085
	.6	.554	.180	.093	.573	.196	.104	.594	.204	.108
Forward-Backward	0	.182	.088	.059	.195	.099	.066	.206	.106	.070
	.3	.303	.117	.069	.321	.132	.081	.337	.140	.085
	.6	.511	.173	.092	.540	.193	.103	.563	.201	.107
Translated	0	.083	.062	.047	.090	.071	.054	.097	.077	.057
	.3	.123	.074	.051	.132	.086	.061	.141	.092	.064
	.6	.258	.119	.071	.274	.135	.081	.290	.142	.084
δ-recursive	0	.083	.067	.053	.090	.077	.060	.097	.083	.064
	.3	.123	.078	.055	.132	.091	.066	.141	.097	.069
	.6	.258	.118	.069	.274	.134	.079	.290	.141	.083

Table 3. Bias of the ten lag-one autocorrelation estimators in series with different lengths. Average: bias averaged across $-.9 \le \rho_1 \le .9$.

Estimators	Auto-correlation	SERIES LENGTH								
		5	10	15	5	10	15	5	10	15
		Exponential errors			Normal errors			Uniform errors		
Conventional	Average	−.220	−.115	−.077	−.224	−.118	−.079	−.225	−.118	−.079
	−.6.	.158	.103	.075	.166	.108	.079	.172	.112	.082
	−.3.	−.018	.001	.003	−.015	.006	.007	−.012	.008	.008
Modified	Average	−.275	−.128	−.082	−.279	−.131	−.084	−.281	−.131	−.085
	−.6.	.048	.048	.037	.058	.054	.041	.065	.058	.045
	−.3.	−.098	−.032	−.018	−.093	−.027	−.014	−.090	−.024	−.013
Cyclic	Average	−.277	−.129	−.082	−.277	−131	−.084	−.278	−.131	−.084
	−.6.	.124	.093	.071	.134	.099	.075	.139	.103	.078
	−.3.	−.063	−.009	−.001	−.057	−.004	.002	−.055	−.001	.004
Exact	Average	−.245	−.119	−.077	−.257	−124	−.080	−.256	−.124	−.080
	−.6.	-.043	.049	.040	.344	.057	.045	.063	.062	.049
	−.3.	−.102	−.035	−.019	−.093	−.027	−.014	−.089	−.023	−.012
C statistic	Average	−.014	−.012	−.010	−.020	−.016	−.012	−.023	−.017	−.013
	−.6.	.339	.193	.133	.350	.196	.137	.350	.200	.139
	−.3.	.172	.096	.067	.174	.100	.070	.177	.103	.072
Fuller	Average	−.012	−.025	−.021	−.016	−.028	−.022	−.018	−.028	−.023
	−.6.	.340	.180	.123	.350	.186	.128	.355	.190	.131
	−.3.	.186	.095	.065	.189	.100	.069	.191	.102	.070
Least Squares	Average	−.276	−.127	−.081	−.281	−.129	−.082	−.281	−.129	−.082
	−.6.	.045	.044	.035	.047	.049	.039	.056	.054	.043
	−.3.	−.095	−.032	−.018	−.094	−.027	−.014	−.092	−.025	−.013
Forward-Backward	Average	−.255	−.120	−.078	−.262	−.128	−.081	−.264	−.125	−.081
	−.6.	.076	.059	.044	.085	.065	.048	.091	.069	.052
	−.3.	−.081	−.027	−.016	−.076	−.023	−.012	−.074	−.020	−.010
Translated	Average	.020	−.015	−.011	−.024	−.018	−.012	−.025	.−.018	−.012
	−.6.	.358	.203	.141	.366	.193	.145	.372	.197	.148
	−.3.	.182	.101	.070	.185	.106	.073	.188	.108	.075
δ-recursive	Average	−.020	−.016	−.011	−.024	−.019	−.013	−.025	−.019	−.013
	−.6.	.358	.187	.114	.366	.208	.118	.372	.212	.121
	−.3.	.182	.093	.056	.185	.098	.060	.188	.101	.061

Table 4. Bias of the ten lag-one autocorrelation estimators in series with different lengths.

Estimators	Auto-correlation	SERIES LENGTH								
		5	10	15	5	10	15	5	10	15
		Exponential errors			Normal errors			Uniform errors		
Conven-tional	0	−.200	−.100	−.066	−.200	−.100	−.067	−.200	−.100	−.066
	.3	−.398	−.208	−.140	−.402	−.213	−.145	−.408	−.217	−.147
	.6	−.615	−.338	−.229	−.629	−.351	−.238	−.640	−.356	−.241
Modified	0	−.250	−.111	−.071	−.250	−.111	−.071	−.250	−.111	−.070
	.3	−.422	−.198	−.128	−.428	−.204	−.134	−.436	−.208	−.136
	.6	−.618	−.309	−.203	−.637	−.324	−.212	−.649	−.329	−.215
Cyclic	0	−.250	−.111	−.071	−.250	−.111	−.072	−.250	−.111	−.070
	.3	−.458	−.220	−.145	−.463	−.226	−.150	−.468	−.229	−.152
	.6	−.697	−.353	−.234	−.703	−.365	−.242	−.711	−.369	−.245
Exact	0	−.235	−.109	−.071	−.238	−.110	−.071	−.239	−.111	−.070
	.3	−.376	−.189	−.125	−.395	−.199	−.132	−.405	−.204	−.135
	.6	−.545	−.292	−.196	−.581	−.311	−.207	−.590	−.317	−.211
C statistic	0	.000	.000	.000	.000	.000	.000	.000	.000	.001
	.3	−.186	−.105	−.072	−.190	−.109	−.077	−.197	−.113	−.079
	.6	−.383	−.226	−.158	−.402	−.240	−.167	−.415	−.245	−.171
Fuller	0	.019	.003	.001	.017	.002	.001	.016	.001	.001
	.3	−.171	−.105	−.073	−.178	−.111	−.079	−.186	−.115	−.081
	.6	−.386	−.242	−.171	−.402	−.255	−.179	−.414	−.260	−.183
Least Squares	0	−.025	−.111	−.071	−.251	−.071	−.071	−.250	−.111	−.071
	.3	−.424	−.197	−.128	−.424	−.201	−.133	−.431	−.205	−.135
	.6	−.613	−.300	−.196	−.625	−.311	−.204	−.635	−.316	−.208
Forward-Backward	0	−.238	−.109	−.070	.240	−.071	−.071	−.241	−.111	−.070
	.3	−.410	−.196	−.128	−.419	−.203	−.134	−.427	−.208	−.136
	.6	−.602	−.304	−.201	−.625	−.320	−.211	−.636	−.326	−.215
Translated	0	.000	.000	.000	.000	.000	.000	.000	.000	.001
	.3	−.198	−.108	−.073	−.202	−.113	−.078	−.208	−.117	−.080
	.6	−.415	−.238	−.163	−.429	−.251	−.171	−.440	−.256	−.174
δ-recursive	0	.000	.000	.000	.000	.000	.000	.000	.000	.001
	.3	−.198	−.101	−.059	−.202	−.106	−.065	−.208	−.110	−.067
	.6	−.415	−.224	−.136	−.429	−.237	−.145	−.440	−.243	−.149

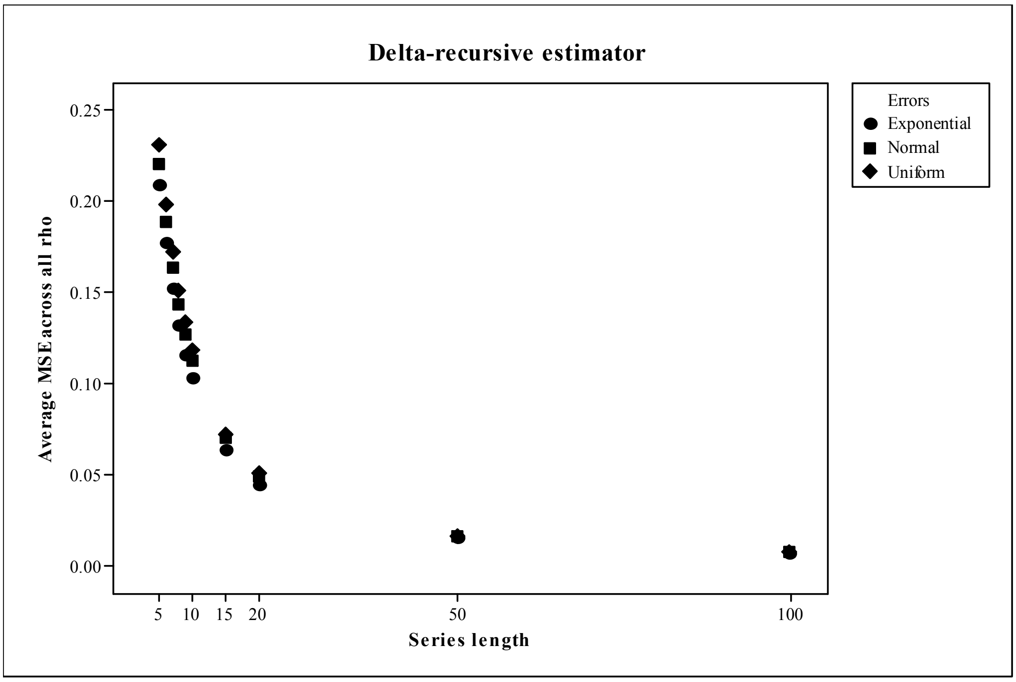

Figure 2. Mean square error (averaged across all ρ_1) for the δ-recursive estimator applied to series with different lengths and errors' distributions.

Monte Carlo sampling: Statistical power

Method

In a first stage the 1% and 5% cut-off points were estimated for each estimator sampling distribution and each series length. In contrast with previous studies (e.g., Huitema & McKean, 1994b; 2000), Monte Carlo methods based on 300,000 iterations were used to estimate the cut-off points, as an alternative to asymptotic tests, as those do not seem to be appropriate for short series (Huitema & McKean, 1991). That is, the power estimates presented here are not founded on a test statistic based on large-sample properties. Instead, the statistical tests associated with the autocorrelation estimators were based on Monte Carlo sampling, which is a suitable approach when the sampling distribution of the test statistic is not known (Noreen, 1989). The analysis was based on nondirectional null hypotheses (H_0: ρ_1 = .0) and, thus, the values corresponding to quantiles .005 and .995 for 1% alpha and quantiles .025 and .975 for 5% alpha were identified. Power was estimated as the proportion of values smaller than the lower bound or greater than the upper bound out of 300,000 iterations per parameter level.

Results

The differences between the best and worst performers in terms of power are generally small, as can be seen comparing the first and the second column of Tables 5, 6, and 7. The proposed estimator performs approximately as the best performers in each condition. In general, sensitivity is rather low in short series and unless the applied researcher has at least 20 measurement times, high degrees of $|\rho|$ may not be reliably detected as statistically significant (Table 7).

If a 1% alpha level is chosen, Type II errors would be excessively frequent for series shorter than 50 observations. Greater power was found for series with exponentially distributed errors – exactly the case for which MSE was lower. Correspondingly, uniform errors' distribution was associated with less sensitivity.

Discussion

The present investigation extends previous research on autocorrelation estimators comparing ten estimators (including a new bias-reducing proposal) in terms of two types of statistical error, bias and variance, summarised as mean square error. Current results concur with previous findings on the existence of bias of autocorrelation estimators applied to short data series, especially in the case of $\rho_1 > 0$, as reported by Matyas and Greenwood (1991). It was also replicated that the *translated* estimator is less bias for positive autocorrelation and more biased for negative one than the conventional estimator (Huitema & McKean, 1991). In general, all estimators studied show lower MSE for negative values of the autoregressive parameter. However, there is not a single optimal estimator for all levels of autocorrelation and all series lengths, as the comparison in terms of MSE values and bias suggests. Bias is present in independent data and gets more pronounced in short autocorrelated series. Out of all of the estimators tested only the *δ-recursive*, the *translated*, and the *C* statistic are not biased for independent series. The magnitude of the bias is heterogeneous among the estimators and, as expected, tends to decrease for longer series. The presence of negative bias when $\rho_1 > 0$ implies that an existing positive serial dependence will be underestimated.

The positive bias in conditions with $\rho_1 < 0$ also entails that the autocorrelation estimate will be closer to zero than it should be. In both cases, it will be harder for the estimates to reach statistical significance when testing H_0: $\rho_1 = 0$.

Table 5. Power estimates for 5% alpha for five-measurement series and several values of the autoregressive parameter. The first column represents the most sensitive test for each error distribution; the second contains the less sensitive one; and third focuses on the proposed estimator.

		Exponential error	
ρ_1	*C* statistic	Circular	δ-recursive
−.6	.1430	.0954	.1348
−.3	.0674	.0599	.0634
.0	.0504	.0501	.0501
.3	.0643	.0559	.0616
.6	.1224	.0683	.0976
		Normal error	
ρ_1	FBackward	Circular	δ-recursive
−.6	.1358	.0859	.1340
−.3	.0656	.0570	.0658
.0	.0507	.0503	.0502
.3	.0630	.0549	.0628
.6	.0934	.0624	.0910
		Uniform error	
ρ_1	*C* statistic	Circular	δ-recursive
−.6	.1175	.0799	.1180
−.3	.0640	.0576	.0626
.0	.0499	.0488	.0504
.3	.0598	.0542	.0603
.6	.0874	.0626	.0788

Table 6. Power estimates for 5% alpha for ten-measurement series and several values of the autoregressive parameter. The first column represents the most sensitive test for each error distribution; the second contains the less sensitive one; and third focuses on the proposed estimator.

	Exponential error		
ρ_1	Translated	*C* statistic	δ-recursive
−.6	.4876	.4463	.4877
−.3	.1528	.1537	.1529
.0	.0502	.0499	.0504
.3	.1076	.0962	.1081
.6	.2991	.2643	.3000
	Normal error		
ρ_1	FBackward	Circular	δ-recursive
−.6	.3803	.3462	.3799
−.3	.1153	.1087	.1177
.0	.0497	.0496	.0497
.3	.1124	.1069	.1126
.6	.2980	.2684	.2913
	Uniform error		
ρ_1	Least Sq	Circular	δ-recursive
−.6	.3477	.3160	.3521
−.3	.1085	.1042	.1128
.0	.0502	.0502	.0501
.3	.1063	.0986	.1037
.6	.2706	.2346	.2574

Table 7. Power estimates for 5% alpha for twenty-measurement series and several values of the autoregressive parameter. The first column represents the most sensitive test for each error distribution; the second contains the less sensitive one; and third focuses on the proposed estimator.

	Exponential error		
ρ_1	FBackward	*C* statistic	δ-recursive
−.6	.8167	.7761	.8095
−.3	.3368	.3215	.3321
.0	.0506	.0501	.0500
.3	.1981	.1803	.1986
.6	.6694	.6345	.6674
	Normal error		
ρ_1	Least Sq	*C* statistic	δ-recursive
−.6	.7287	.6993	.7242
−.3	.2307	.2251	.2308
.0	.0488	.0491	.0485
.3	.2262	.2182	.2253
.6	.6677	.6575	.6606
	Uniform error		
ρ_1	Least Sq	Circular	δ-recursive
−.6	.7080	.6828	.7061
−.3	.2210	.2112	.2227
.0	.0505	.0505	.0507
.3	.2145	.2048	.2135
.6	.6431	.6186	.6369

The variance of the estimators is also dissimilar and the efficiency of the estimators depends on the autoregressive parameter and series length. Therefore, there is not a single uniform minimum variance unbiased estimator among the ones assessed in the present study. The proposed *δ-recursive* estimator equals or improves the performance of the other estimators (in terms of MSE and bias) when $n \geq 10$ in the cases of positive autocorrelation and considering the overall performance across all ρ_1. Therefore, it can be considered a viable alternative whenever the sign of the autoregressive parameter is not known or is supposed to be positive. For series with less than ten measurement times, the *Fuller* and the *translated* estimators are the most adequate ones if the applied researcher assumes that $\rho_1 \geq 0$ or has no information about the possible direction of the serial dependence. For $\rho_1 < 0$ the *conventional* estimator is the one showing best results for all series lengths studied.

The present study also estimates power using tests based on Monte Carlo sampling rather than on asymptotic formulae, as has been previously done. The estimates obtained here are somewhat higher than the ones reported for Bartlett's test (Arnau & Bono, 2001; Huitema & McKean, 1991) and somewhat lower than the ones associated with the test recommended by Huitema and McKean (1991). Regarding Moran's (1948) approximation for the *conventional* estimator, the Monte Carlo sampling tests are more sensitive for $\rho_1 > 0$ and less sensitive for $\rho_1 < 0$. For the *translated* estimator, power estimates are similar for Monte Carlo sampling and Moran's approximation (Arnau & Bono, 2001). In general, present and past findings coincide in the low sensitivity in short data series. The difference in power between the tests associated with the estimators is only slight.

Combining the findings of previous research and the present investigation it seems that empirical studies on real behavioural measurements (e.g., the surveys by Busk and Marascuilo, 1988; Huitema, 1985; and Parker, 2006) are not likely to resolve unequivocally the question of the existence and statistical significance of serial dependence in single-case data. The reason is the high statistical error of the estimators applied to short data series and the lack of power of the test associated with those estimators. Only for series containing 50 or 100 data points would the evidence have any meaning.

For applied researchers the lack of precision and sensitivity in estimating autocorrelation implies uncertainty about the degree of serial dependence that may be present in the behavioural data collected. It has been remarked that low estimates of serial dependence do not guarantee the adequacy of applying statistical techniques based on the General Linear Model to assess intervention effectiveness (Ferron, 2002). Therefore, clinical, educational, and social psychologists need to assess intervention effectiveness by means of procedures with appropriate Type I and Type II error rates in presence of autocorrelation.

A specific contribution of the present study to methodological research is the comparison between errors' distribution shapes. The results indicate that generating data with errors following a normal, a rectangular or a highly asymmetric distribution does not influence critically the MSE and power estimates. Hence, the findings of studies based solely on normally distributed errors may not be limited to the conditions actually simulated.

A limitation of the present study consists in the fact that only an AR(1) model was employed to generate data. As it has been pointed out (Harrop & Velicer, 1985), there are other models that may be used to represent behavioural data. Future studies may be based, for instance, on moving average models to extend the evidence on the performance of autocorrelation estimators. Additionally, in view of the presence of bias in each successive estimator proposed by different authors, a bias reducing technique may be useful. The bootstrap adjustment of bias has been shown to be effective correcting the positive bias for $\rho_1 < 0$ and the negative one for $\rho_1 > 0$ and reducing the MSE, according to the data presented by McKnight et al. (2000) for series with $n \geq 20$, in contrast to jackknife methods which increase the error variance (Huitema & McKean, 1994a). We consider that bootstrap ought to be applied to the estimators that seem to have a more adequate performance in terms of MSE – the *Fuller*, the *translated*, and the *δ-recursive* estimators when positive serial dependence is assumed or when the sign of the autocorrelation is unknown, and the *conventional* estimator for negative one. Therefore, it is necessary to investigate the degree to which the bootstrap improves those estimators when few measurements are available, as is the case in applied psychological studies. Another possible application of the bootstrap is to construct confidence intervals about the autocorrelation estimates, since those have shown appropriate coverage (McKnight et al., 2000), and use them to make statistical decisions. Bootstrap has the advantage of allowing asymmetric confidence intervals which correspond to the skewed distributions of the estimators for short data series. In this case, the power of the tests based on bootstrap confidence intervals has to be compared to the sensitivity of the test constructed using Monte Carlo sampling, since Bartlett's (1946) and Moran's (1948) approximations for hypothesis testing seem inappropriate for short data series (Arnau & Bono, 2001; Huitema & McKean, 1991; Matyas & Greenwood, 1991).

RESUMEN

Autocorrelación de primer orden en series cortas: Estimación y prueba de hipótesis. La primera parte del estudio consiste en revisar nueve estimadores del parámetro autorregresivo de primer orden y proponer un estimador nuevo. Las relaciones y diferencias entre los estimadores se explican para conseguir una diferenciación mejor entre ellos. En la segunda parte del estudio se explora la precisión de la estimación de la autocorrelación. El rendimiento de los diez estimadores se compara en términos de error cuadrático medio, combinando sesgo y varianza, utilizando series de datos generadas mediante simulación Monte Carlo. Los resultados muestran que no hay un estimador óptimo para todas las condiciones, sugiriendo que el estimador a utilizar debería escogerse según la longitud de las series y la información disponible sobre la posible dirección de la dependencia serial. Además, la probabilidad de etiquetar una autocorrelación existente como estadísticamente significativa se estudió mediante muestreo Monte Carlo. Las pruebas asociadas con los diferentes estimadores muestran potencia similar, observándose que es poco probable detectar la dependencia serial si se dispone de menos de 20 medidas.

REFERENCES

Anderson, R. L. (1942). Distribution of the serial correlation coefficient. *The Annals of Mathematical Statistics, 13*, 1-13.

Arnau, J. (1999). Reducción del sesgo en la estimación de la autocorrelación en series temporales cortas. *Metodología de las Ciencias del Comportamiento, 1*, 25-37.

Arnau, J., & Bono, R. (2001). Autocorrelation and bias in short time series: An alternative estimator. *Quality & Quantity, 35*, 365-387.

Bartlett, M. S. (1946). On the theoretical specification and sampling properties of autocorrelated time-series. *Journal of the Royal Statistical Society, 8*, 27-41.

Box, G. E. P., & Jenkins, G. M. (1970). *Time series analysis, forecasting and control*. San Francisco: Holden-Day.

Bradley, J. V. (1977). A common situation conducive to bizarre distribution shapes. *American Statistician, 31*, 147-150.

Busk, P. L., & Marascuilo, L. A. (1988). Autocorrelation in single-subject research: A counterargument to the myth of no autocorrelation. *Behavioral Assessment, 10*, 229-242.

Cox, D. R. (1966). The null distribution of the first serial correlation coefficient. *Biometrika, 43*, 623-626.

Crosbie, J. (1989). The inapproprieteness of the C Statistic for assessing stability or treatment effects with single-subject data. *Behavior Assessment, 11*, 315-325.

DeCarlo, L. T., & Tryon, W. W. (1993). Estimating and testing correlation with small samples: A comparison of the C-statistic to modified estimator. *Behaviour Research and Therapy, 31*, 781-788.

Edgington, E. S., & Onghena, P. (2007). *Randomization tests* (4th ed.). London: Chapman & Hall/CRC.

Ferron, J. (2002). Reconsidering the use of the general linear model with single-case data. *Behavior Research Methods, Instruments, & Computers, 34*, 324-331.

Ferron, J., & Ware, W. (1995). Analyzing single-case data: The power of randomization tests. *The Journal of Experimental Education, 63*, 167-178.

Fuller, W. A. (1976). *Introduction to statistical time series.* New York: John Wiley & Sons.

Greenwood, K. M., & Matyas, T. A. (1990). Problems with application of interrupted time series analysis for brief single-subject data. *Behavioral Assessment, 12*, 355-370.

Harrop, J. W., & Velicer, W. F. (1985). A comparison of alternative approaches to the analysis of interrupted time-series. *Multivariate Behavioral Research, 20*, 27-44.

Huitema, B. E. (1985). Autocorrelation in applied behavior analysis: A myth. *Behavior Assessment, 7*, 107-118.

Huitema, B. E. (1988). Autocorrelation: 10 years of confusion. *Behavioral Assessment, 10*, 253-294.

Huitema, B. E., & McKean, J. W. (1991). Autocorrelation estimation and inference with small samples. *Psychological Bulletin, 110,* 291-304.

Huitema, B. E., & McKean, J. W. (1994a). Reduced bias autocorrelation estimation: Three jackknife methods. *Educational and Psychological Measurement, 54*, 654-665.

Huitema, B. E., & McKean, J. W. (1994b). Tests of H_0: $\rho_1 = 0$ for autocorrelation estimators r_{F1} and r_{F2}. *Perceptual and Motor Skills, 78*, 331-336.

Huitema, B. E., & McKean, J. W. (1994c). Two reduced-bias autocorrelation estimators: r_{F1} and r_{F2}. *Perceptual and Motor Skills, 78*, 323-330.

Huitema, B. E., & McKean, J. W. (2000). A simple and powerful test for autocorrelated errors in OLS intervention models. *Psychological Reports, 87*, 3-20.

Huitema, B. E., & McKean, J. W. (2007a). An improved portmanteau test for autocorrelated errors in interrupted time-series regression models. *Behavior Research Methods, 39*, 343-349.

Huitema, B. E., & McKean, J. W. (2007b). Identifying autocorrelation generated by various error processes in interrupted time-series progression designs: A comparison of AR1 and portmanteau tests. *Educational and Psychological Measurement, 67*, 447-459.

Jenkins, G. M., & Watts, D. G. (1968). *Spectral analysis and its applications.* New York: Holden-Day.

Kendall, M. G. (1954). Note on bias in the estimation of estimation of autocorrelation. *Biometrika, 41*, 403-404.

Kendall, M. G., & Ord, J. K. (1990). *Time series.* Sevenoaks: Edward Arnold.

Manolov, R., & Solanas, A. (2009). Problems of the randomization test for AB designs. *Psicológica, 30,* 137-154.

Marriott, F. H. C., & Pope, A. (1954). Bias in the estimation of autocorrelation. *Biometrika, 41*, 390-402.

Matyas, T. A., & Greenwood, K. M. (1991). Problems in the estimation of autocorrelation in brief time series and some implications for behavioral data. *Behavior Assessment, 13*, 137-157.

Matyas, T. A., & Greenwood, K. M. (1997). Serial dependency in single-case time series. In R. D. Franklin, D. B. Allison, & B. S. Gorman (Eds.), *Design and analysis of single-case research* (pp. 215-243). Mahwah, NJ: Lawrence Erlbaum.

McKean, J. W., & Huitema, B. E. (1993). Small sample properties of the Spearman autocorrelation estimator. *Perceptual & Motor Skills, 76*, 384-386.

McKnight, S. D., McKean, J. W., & Huitema, B. E. (2000). A double bootstrap method to analyze linear models with autoregressive error terms. *Psychological Methods, 3*, 87-101.

Micceri, T. (1989). The unicorn, the normal curve, and other improbable creatures. *Psychological Bulletin, 105*, 156-166.

Moran, P. A. P. (1948). Some theorems on time series. II. The significance of the serial correlation coefficient. *Biometrika, 35*, 255-267.

Moran, P. A. P. (1967). Testing for serial correlation with exponentially distributed variates. *Biometrika, 54*, 395-401.

Moran, P. A. P. (1970). A note on serial correlation coefficients. *Biometrika, 57*, 670-673.

Noreen, E. R. (1989). *Computer-intensive methods for testing hypotheses: An introduction.* New York: John Wiley and Sons.

Orcutt, G. H. (1948). A study of the autoregressive nature of the time series used for Tinbergen's model of the economic system of the United States, 1919-1932. *Journal of the Royal Statistical Society (Series B), 10*, 1-53.

Parker, R. I. (2006). Increased reliability for single-case research results: Is bootstrap the answer? *Behavior Therapy, 37*, 326-338.

Quenouille, M. (1949). Approximate test of correlation in time series. *Journal of the Royal Statistical Society (Series B), 11*, 18-84.

Sawilowsky, S. S., & Blair, R. C. (1992). A more realistic look at the robustness and Type II error properties of the *t* test to departures from population normality. *Psychological Bulletin, 111*, 352-360.

Scheffé, H. (1959). *The analysis of variance.* New York: Wiley.

Sharpley, C. F., & Alavosius, M. P. (1988). Autocorrelation in behavior data: An alternative perspective. *Behavior Assessment, 10*, 243-251.

Simonton, D. K. (1977). Cross-sectional time-series experiments: Some suggested statistical analyses. *Psychological Bulletin, 84*, 489-502.

Spanos, A. (1987). *Statistical foundations of econometrics modeling.* Cambridge: Cambridge University Press.

Suen, H. K., & Ary, D. (1987). Autocorrelation in behavior analysis: Myth or reality? *Behavioral Assessment, 9*, 150-130.

Toothaker, L. E., Banz, M., Noble, C., Camp, J., & Davis, D. (1983). N = 1 designs: The failure of ANOVA-based tests. *Journal of Educational Statistics, 4*, 289-309.

Tryon, W. W. (1982). A simplified time-series analysis for evaluating treatment interventions. *Journal of Applied Behavior Analysis, 15*, 423-429.

Tryon, W. W. (1984). "A simplified time-series analysis for evaluating treatment interventions": A rejoinder to Blumberg. *Journal of Applied Behavior Analysis, 17*, 543-544.

Tuan, D. P. (1992). Approximate distribution of parameter estimators for first-order autoregressive models. *Journal of Time Series Analysis, 13*, 147-170.

Velicer, W. F., & Harrop, J. W. (1983). The reliability and accuracy of the time-series model identification. *Evaluation Review, 7*, 551-560.

Weisberg, S. (1980). *Applied linear regression.* New York: John Wiley & Sons.

Young, L. C. (1941). On the randomness in ordered sequences. *The Annals of Mathematical Statistics, 12*, 293-300.

(Manuscript received: 8 May 2009; accepted: 25 September 2009)