Estimation of the mean and autocorrelations of a stationary process [*]

Jean-Marie Dufour [†]

First version: December 1998
This version: May 3, 2015, 5:51pm

[*] This work was supported by the Canada Research Chair Program (Chair in Econometrics, Université de Montréal), the Canadian Network of Centres of Excellence [program on *Mathematics of Information Technology and Complex Systems* (MITACS)], the Canada Council for the Arts (Killam Fellowship), the Natural Sciences and Engineering Research Council of Canada, the Social Sciences and Humanities Research Council of Canada, and the Fonds FCAR (Government of Québec).

[†] Canada Research Chair Holder (Econometrics). Centre de recherche et développement en économique (C.R.D.E.), Centre interuniversitaire de recherche en analyse des organisations (CIRANO), and Département de sciences économiques, Université de Montréal. Mailing address: Département de sciences économiques, Université de Montréal, C.P. 6128 succursale Centre-ville, Montréal, Québec, Canada H3C 3J7. TEL: 1 514 343 2400; FAX: 1 514 343 5831; e-mail: jean.marie.dufour@umontreal.ca. Web page: http://www.fas.umontreal.ca/SCECO/Dufour .

Contents

List of Definitions, Assumptions, Propositions and Theorems

1. General distributional results

1.1 Suppose we have T observations $X_1, X_2, \ldots, X_T$ from a realization of a second-order stationary process. The natural estimators of the first and second moments of the process are: for the mean,

$$\overline{X}_T = \frac{1}{T}\sum_{t=1}^{T} X_t \ ,$$

for the autocovariances

$$c_k = \frac{1}{T}\sum_{t=1}^{T-k}\left(X_t - \overline{X}_T\right)\left(X_{t+k} - \overline{X}_T\right) \ , \quad 1 \le k \le T-1 \ ,$$

and for the autocorrelations

$$r_k = c_k / c_0 \ , \quad 1 \le k \le T-1 \ .$$

1.2 Theorem DISTRIBUTION OF THE ARITHMETIC MEAN. *Let* $\{X_t : t \in \mathbb{Z}\}$ *a second-order stationary process with mean* μ, *and let* $\overline{X}_T = \sum_{t=1}^{T} X_t / \, T$. *Then*

(1) $E\left(\overline{X}_T\right) = \mu$ *and* $\overline{X}_T$ *is an unbiased estimator of* μ;

(2) $Var\left(\overline{X}_T\right) = \frac{1}{T}\sum_{k=-(T-1)}^{T-1}\left(1 - \frac{|k|}{T}\right)\gamma_x(k)$;

(3) *if* $\gamma_x(k) \underset{k\to\infty}{\longrightarrow} 0$,

$$Var\left(\overline{X}_T\right) \underset{T\to\infty}{\longrightarrow} 0 \text{ and } \overline{X}_T \underset{T\to\infty}{\overset{m.q.}{\longrightarrow}} \mu \ ;$$

(4) *if the series* $\sum_{k=-\infty}^{\infty}\gamma_x(k)$ *converges, then*

$$\lim_{T\to\infty} T \, Var\left(\overline{X}_T\right) = \sum_{k=-\infty}^{\infty}\gamma_x(k) \ ;$$

(5) *if the spectral density* $f_x(\omega)$ *exists and is continuous at* $\omega = 0$ *, then*

$$\lim_{T\to\infty} T\, Var\left(\overline{X}_T\right) = 2\pi f_x(0)\ ;$$

(6) *if*

$$X_t = \mu + \sum_{j=-\infty}^{\infty} \psi_j u_{t-j}\,,\ where\ \{u_t : t \in \mathbb{Z}\} \sim IID\left(0, \sigma^2\right)\,,$$

and

$$\sum_{j=-\infty}^{\infty} |\psi_j| < \infty\,,$$

then

$$\sqrt{T}\left(\overline{X}_T - \mu\right) \xrightarrow[T\to\infty]{L} N\left[0, \sum_{k=-\infty}^{\infty} \gamma_x(k)\right]$$

and

$$\sum_{k=-\infty}^{\infty} \gamma_x(k) = \sigma^2 \left(\sum_{j=-\infty}^{\infty} \psi_j\right)^2 .$$

PROOF. See Anderson (1971, Sections 8.3.1 and 8.4.1) and Brockwell and Davis (1991, Section 7.1). □

1.3 Theorem DISTRIBUTION OF SAMPLE AUTOCORRELATIONS FOR A LINEAR STATIONARY PROCESS. *Let* $X_t = \mu + \sum_{j=-\infty}^{\infty} \psi_j u_{t-j}$*, where* $\sum_{j=-\infty}^{\infty} |\psi_j| < \infty$ *and* $\{u_t : t \in \mathbb{Z}\} \sim IID\left(0, \sigma^2\right)$*. If*

$$(a)\ \sum_{j=-\infty}^{\infty} |j|\psi_j^2 < \infty$$

or

$$(b)\ E\left(u_t^4\right) < \infty,\ \forall t\,,$$

then the asymptotic distribution of the vector

$$\left[\sqrt{T}\left(r_1-\rho_1\right),\sqrt{T}\left(r_2-\rho_2\right),\ldots,\sqrt{T}\left(r_m-\rho_m\right)\right]'$$

is $N[0,W_m]$ *as* $T\to\infty$, *where* $\rho_k=\gamma_x(k)/\gamma_x(0)$, $W_m=[w_{jk}]_{j,\,k=1,\ldots,\,m}$ *and*

$$\begin{aligned}
w_{jk} &= \sum_{h=-\infty}^{\infty}\left(\rho_{h+j}\rho_{h+k}+\rho_{h-j}\rho_{h+k}-2\,\rho_k\rho_h\rho_{h+j}-2\,\rho_j\rho_h\rho_{h+k}+2\,\rho_j\rho_k\rho_h^2\right)\\
&= \sum_{h=1}^{\infty}\left(\rho_{h+j}+\rho_{h-j}-2\,\rho_j\rho_h\right)\left(\rho_{h+k}+\rho_{h-k}-2\,\rho_k\rho_h\right)\\
&= \frac{4\pi}{\gamma_x(0)^2}\int_{-\pi}^{\pi}\left[\cos(\omega j)-\rho_j\right]\left[\cos(\omega k)-\rho_k\right]f_x(\omega)^2\,d\omega\,.
\end{aligned}$$

PROOF. See Anderson (1971, Theorem 8.4.6, p. 489) and Brockwell and Davis (1991, Theorems 7.2.1 and 7.2.2). □

1.4 The expressions w_{jk} are called Bartlett's formula for the covariances of the autocorrelations. The formula w_{jk} may also be written

$$\begin{aligned}
w_{jk} &= \left(\lambda_{j+k}+\lambda_{j-k}-2\rho j\lambda_k-2\rho_k\lambda_j+2\rho_j\rho_k\lambda_0\right)/\gamma_0^2\\
&= \bar{\lambda}_{j+k}+\bar{\lambda}_{j-k}-2\rho_j\bar{\lambda}_k-2\rho_k\bar{\lambda}_j-2\rho_j\rho_k\bar{\lambda}_0
\end{aligned}$$

where

$$\lambda_i=\sum_{h=-\infty}^{\infty}\gamma_h\gamma_{h+i}\,,\quad \bar{\lambda}_i\equiv\lambda_i/\gamma_0^2=\sum_{h=-\infty}^{\infty}\rho_h\rho_{h+i}\,.$$

2. Special cases

2.1 ASYMPTOTIC VARIANCE. Under the conditions of Theorem **1.3**, the asymptotic distribution of $\sqrt{T}\,(r_k - \rho_k)$ is $N\,[0, w_{kk}]$, where

$$\begin{aligned} w_{kk} &= \sum_{h=-\infty}^{\infty} \left(\rho_{h+k}^2 + \rho_{h-k}\rho_{h+k} - 4\,\rho_k\rho_h\rho_{h+k} + 2\,\rho_k^2\rho_h^2\right) \\ &= \sum_{h=-\infty}^{\infty} \left(\rho_h^2 + \rho_h\rho_{h+2k} - 4\,\rho_h\rho_k\rho_{h+k} + 2\,\rho_h^2\rho_k^2\right) \\ &= \sum_{h=1}^{\infty} \left(\rho_{h+k} + \rho_{h-k} - 2\,\rho_h\rho_k\right)^2 . \end{aligned}$$

For T large, $\sqrt{T}\,(r_k - \rho_k) \overset{a}{\sim} N\,[0, w_{kk}]$.

2.2 WHITE NOISE. If

$$\begin{aligned} \rho_k &= 1\ , \text{ for } k = 0\ , \\ &= 0\ , \text{ for } k \neq 0\ , \end{aligned}$$

we find

$$\begin{aligned} w_{jk} &= 1\ , \text{ if } j = k \\ &= 0\ , \text{ if } j \neq k\ . \end{aligned}$$

For T large, the sampling autocorrelations are mutually uncorrelated and

$$\sqrt{T} r_k \overset{a}{\sim} N\,[0, 1]\ , \text{ for } k \geq 1\ .$$

2.3 MA(q) PROCESS. If $\rho_k = 0$, for $|k| \geq q+1$, we find

$$w_{jk} = \sum_{h=1}^{\infty} \rho_{h-j}\rho_{h-k} = \sum_{h=1}^{\infty} \rho_{j-h}\rho_{k-h} = \sum_{h=1}^{\infty} \rho_{k-h+(j-k)}\rho_{k-h}$$

$$= \sum_{h=-\infty}^{k-1} \rho_h \rho_{h+(j-k)} = \sum_{h=-q}^{q-(j-k)} \rho_h \rho_{h+(j-k)}\,, \text{ for } j \geq k \geq q+1\,,$$

hence

$$\begin{aligned} w_{jk} &= 0\,, && \text{if } k \geq q+1 \text{ and } j \geq k+2q+1 \\ &= \textstyle\sum_{h=-q}^{q-(j-k)} \rho_h \rho_{h+(j-k)}\,, && \text{if } q+1 \leq k \leq j \leq k+2q\,. \end{aligned} \tag{2.1}$$

In particular,

$$w_{kk} = \sum_{h=-q}^{q} \rho_h^2 = 1 + 2\sum_{h=1}^{q} \rho_h^2\,, \text{ if } k \geq q+1\,.$$

3. Exact tests of randomness

3.1 Theorem EXACT MOMENTS OF AUTOCORRELATIONS FOR AN *i.i.d.* SAMPLE. *Let the random variables $X_1, \ldots, X_T$ be independent and identically distributed $(i.i.d.)$ according to a continuous distribution. Then*

$$E(r_k) = -\frac{T-k}{T(T-1)}\,, \textit{ for } 1 \leq k \leq T-1\,,$$

and

$$Var(r_k) \leq \overline{V}_k\,,$$

where

$$\overline{V}_k \equiv \frac{T^4 - (k+7)T^3 + (7k+16)T^2 + 2\left(k^2 - 9k - 6\right)T - 4k(k-4)}{T(T-1)^2(T-2)(T-3)}$$

if $1 \leq k < T/2$ *and* $T > 3$, *and*

$$\overline{V}_k \equiv \frac{(T-k)\left[T^2 - 3T - 2(k-2)\right]}{T(T-1)^2(T-3)}$$

if $T/2 \leq k < T$ *and* $T > 3$.

PROOF. See Dufour and Roy (1985). $\square$

3.2 For $k = 1$, we find

$$E(r_1) = -1/T \ ,$$

$$Var(r_1) \leq \frac{T-2}{T(T-1)} \ .$$

By Chebyshev's inequality,

$$P[|r_k - E(r_k)| \geq \lambda] \leq \frac{Var(r_k)}{\lambda^2} \leq \frac{\overline{V}_k}{\lambda^2} \ .$$

3.3 Theorem EXACT MOMENTS OF AUTOCORRELATIONS FOR A GAUSSIAN *i.i.d.* SAMPLE. *Let $X_1, \ldots, X_T$ be i.i.d. random variables following a distribution $N\left[\mu, \sigma^2\right]$ distribution. Then*

$$E(r_k) = -\frac{(T-k)}{T(T-1)} \ , \ for \ 1 \leq k \leq T-1 \ ,$$

$$Var(r_k) = \frac{T^4 - (k+3)T^3 + 3kT^2 + 2k(k+1)T - 4k^2}{(T+1)T^2(T-1)^2}$$

for $1 \leq k < T/2$ *and* $T > 3$, *and*

$$Var(r_k) = \frac{(T-k)(T-2)\left(T^2 + T - 2k\right)}{(T+1)T^2(T-1)^2}$$

for $T/2 \leq k < T$ *and* $T > 3$. *Furthermore, for* $1 \leq k < h \leq T-1$,

$$Cov(r_k, r_h) = \frac{2\left[kh(T-1) - (T-h)\left(T^2 - k\right)\right]}{(T+1)T^2(T-1)^2}$$

if $l < h+k < T$, *and*

$$Cov(r_k, r_h) = \frac{2(T-h)[2k-(k+1)T]}{(T+1)T^2(T-1)^2}$$

if $h+k \geq T$.

PROOF. See Dufour and Roy (1985). □

3.4 For T large, we have

$$\frac{r_k - E(r_k)}{[Var(r_k)]^{1/2}} \overset{a}{\sim} N(0,1) \ .$$

In small or moderately large samples, the normal approximation is much more accurate when the formulae for $E(r_k)$ et $Var(r_k)$ given by Theorem **3.3** are used, rather than $E(r_k) = 0$ et $Var(r_k) = 1/T$; see Dufour and Roy (1985).

References

Anderson, T. W. (1971), *The Statistical Analysis of Time Series*, John Wiley & Sons, New York.

Brockwell, P. J. and Davis, R. A. (1991), *Time Series: Theory and Methods*, second edn, Springer-Verlag, New York.

Dufour, J.-M. and Roy, R. (1985), 'Some exact results on sample autocorrelations and tests of randomness', *Journal of Econometrics* **29**, 257–273. Corrigendum 41(1989), 279-281.