

The Effects of Instantaneous Transformations on the Fractional Differencing Parameter¹

Effetti delle Trasformazioni sulla Stima della Differenza Frazionaria

Angela D'Elia, Domenico Piccolo

Dipartimento di Scienze Statistiche, Università degli Studi di Napoli Federico II

Via L. Rodinò 22, 80138 Napoli

angdelia@unina.it; dopiccol@unina.it

Riassunto: Questo lavoro studia gli effetti del parametro λ della trasformazione Box-Cox sulla stima del parametro d (ordine della differenza frazionaria) in serie generate da processi a lunga memoria. Infatti, le evidenze empiriche mostrano che tali effetti sono notevoli sia a fini di identificazione che nelle procedure inferenziali. A tal fine, l'articolo propone una stima congiunta dei parametri (d, λ) mediante una formulazione normalizzata della verosimiglianza spettrale. Uno studio Monte Carlo conferma il buon comportamento delle stime ottenute in campioni di dimensione finita.

Keywords : Box-Cox Transformation, Fractional Difference, Spectral Likelihood.

1. Introduction

The modifications of the observations, such as the Box-Cox transformation, is a common practice in time series analysis in order to find a convenient metric for the series to be a realization of a Gaussian, homoschedastic, linear process (for recent developments, see: Xia *et al.*, 2000). Of course, a different non-linear scale induces significant modifications on the autocorrelation structure of the process, as discussed by many authors (e.g. Granger and Newbold, 1976).

On the other hand, in applied research (Hidrology, Environmental sciences, Economics, etc.), the “long memory” behaviour of the autocorrelation function is the starting point to detect and estimate the persistence of dependence between observations that are far away. In this regard, different statistical models (Beran, 1994) and, in particular, the class of ARFIMA models has been successfully introduced (Granger and Joyeaux, 1980; Hosking, 1981).

The aim of our paper is to study the effects of transforming an observed series on the estimation of the fractional differencing parameter of a long memory process.

The paper is organized as follows: in Section 2 empirical evidences are shown in order to highlight the effects on the fractional difference parameter of undue transformations of the data; then, in Section 3, we propose a spectral version of the likelihood function for the joint estimation of the parameters (d, λ) . Section 4 is devoted to a Monte Carlo

¹ The work has been supported by research projects of the Dipartimento di Scienze Statistiche, University of Naples Federico II and by the research structures of the Centro per la Formazione in Economia e Politica dello Sviluppo Rurale, Portici. A special thank is due to prof. M. Corduas for encouragements and suggestions in this research. This paper is the result of a joint work; however, A. D'Elia wrote sections 2 and 3, while D. Piccolo wrote sections 1 and 4.

analysis which confirms the usefulness of the approach in finite samples. This final section includes some further comments for future developments.

2. The effect of a “wrong” transformation

A Monte Carlo study has been implemented for exploring how the estimate of the fractional differencing parameter deviates from the true value d of the process, after the application of a transformation of the observed series. In this regard, it is important to notice that, since the true process is assumed known *a priori*, the study highlights the effect on the estimation of d when a “wrong” transformation is applied on the series.

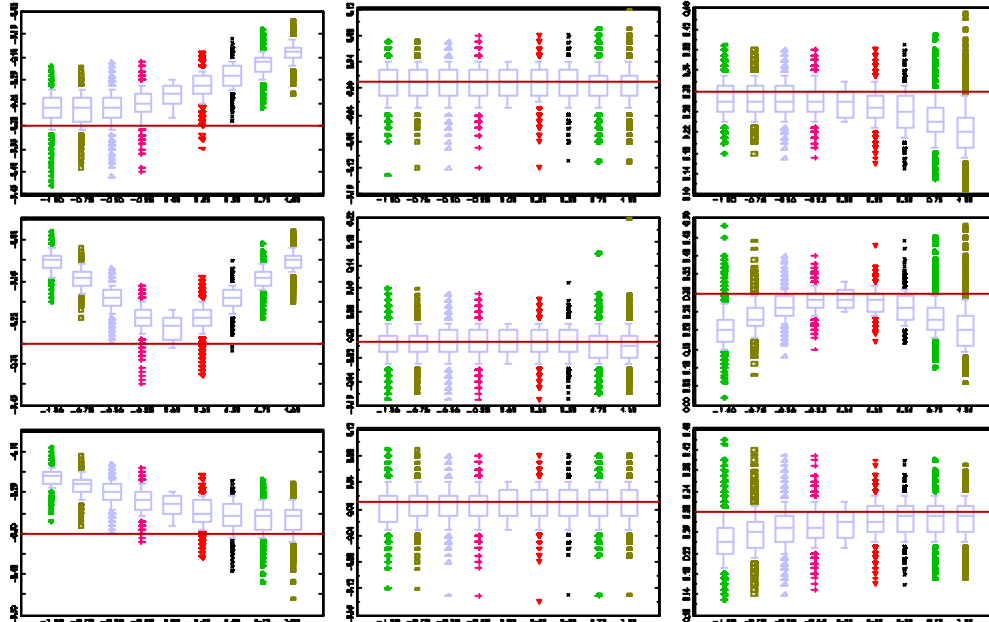
As far as concerns the simulation plan, for different values of d , one thousand series of $n=1000$ observations were generated from the ARFIMA(0, d ,0) process: $\tilde{N}^d X_t = a_t$ where a_t is a standard Gaussian White Noise, and $X_t = W_t^{\lambda_0}$ with a fixed λ_0 . Then, the simulated series were submitted to power transformations for varying λ , by means of the Box-Cox transformation defined by:

$$Y_t(\lambda) = \begin{cases} [(X_t^\lambda - 1) / \lambda], & \text{if } \lambda \neq 0; \\ \ln(X_t), & \text{if } \lambda = 0. \end{cases}$$

On each transformed series the parameter d was estimated using the preliminary consistent method proposed by Corduas (2000).

Figure 1 shows for processes with $\lambda = -0.75, 0, 1$ (1st, 2nd and 3rd row, respectively) and $d = -0.3, 0, 0.3$ (1st, 2nd and 3rd column, respectively) the behaviour of the Monte Carlo estimates of the parameter d . Each window shows the effect of an undue transformation of the series, for some values of $\lambda = -1, -0.75, \dots, 1$; thus, for instance, the central box-plot in each window displays the effect of the logarithmic transformation.

Figure 1: *Effects on the estimation of d , for different transformations of data.*



It seems evident that a wrong transformation of the series produces a significant bias in the estimation of d . In fact, when the true d is positive, an undue transformation makes

the estimate of d negatively biased, while, when d is negative the estimates appear positively biased. No relevant effect seems to happen for $d \cong 0$. Then, these evidences strongly motivate the usefulness of a joint estimation of (λ, d) .

3. The spectral likelihood of the parameters λ and d

Let X_t be an ARFIMA(0, d , 0) process where $d \in (-0.5, 0.5)$; thus, X_t is both stationary and invertible. From a theoretical point of view, the difficulty of studying the relationship between λ and d arises from the definition of a stochastic process as a measurable function on the space $\Omega \times T$, where Ω is the probability space of the events and T is the time parameter space. Clearly, any transformation of X_t affects the random variable $X(\omega, t)$ and its range defined over Ω , while spectral, autocorrelation and modelling analyses are all based on the linear spanning of this random variable over T . Thus, it becomes important to obtain simultaneous estimates of the parameters d and λ , by means of maximum likelihood (ML) methods. Our proposal is to exploit Whittle's approximation to the Gaussian likelihood and to normalize the Box-Cox transformation, in order to make it comparable in the different metrics induced by λ .

Following Brockwell and Davis (1991, pp. 528-529), the log-likelihood function for an ARIMA(0, d , 0) model –concentrating out the σ^2 parameter- is apart from an additive constant:

$$l(d, \lambda; \underline{x}) = \log \left(\frac{1}{n} \sum \frac{I_n(\omega_j; \lambda)}{g(\omega_j)} \right) + \frac{1}{n} \sum \log g(\omega_j),$$

where the sums and the Fourier frequencies $\omega_j = 2\pi j/n$ are defined for $j=1, 2, \dots, [n/2]$, and

$$I_n(\omega_j; \lambda) = \frac{1}{n} \left| \sum_{t=1}^n y_t^*(\lambda) e^{-it\omega_j} \right|^2, \quad g(\omega_j) = \left| 2 \sin \left(\frac{\omega_j}{2} \right) \right|^{-2d}.$$

In the modified periodogram $I_n(\omega_j; \lambda)$ we have defined (as in Brubacher, 1976):

$$\begin{aligned} y_t^*(\lambda) &= [(x_t^\lambda - 1)/(\lambda G_n^{\lambda-1})], & \text{if } \lambda \neq 0, \\ &= G_n \ln(X_t), & \text{if } \lambda = 0, \end{aligned}$$

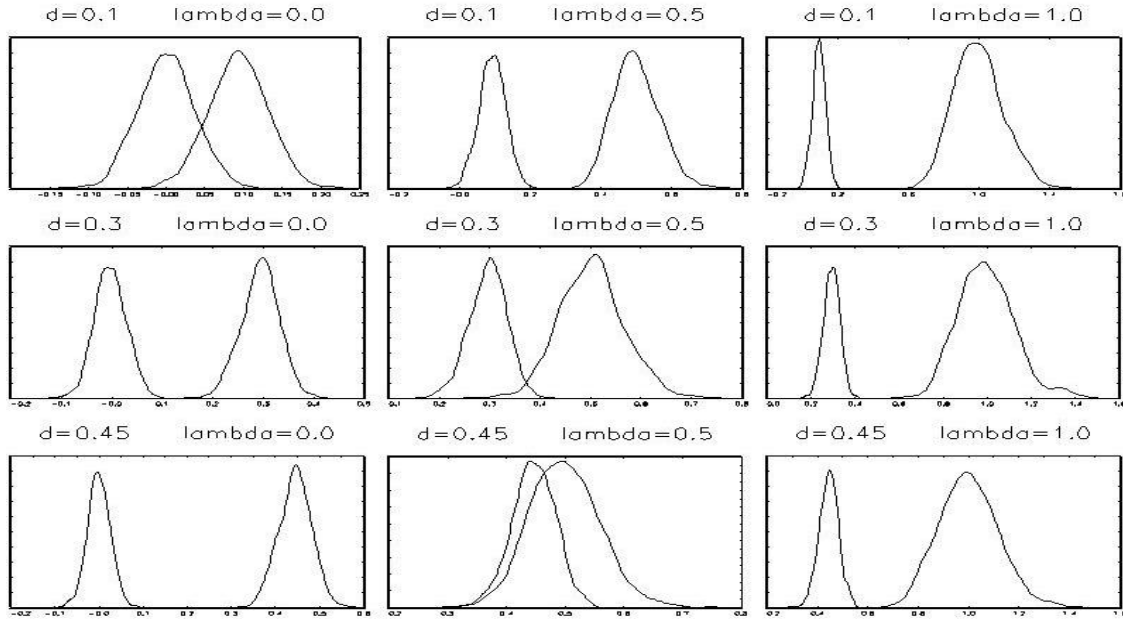
where $G_n = \exp(n^{-1} \sum \log(x_t))$ is the sample geometric mean of the series.

4. A simulation study

In this section we show the main results we obtained from a Monte Carlo study on the behaviour of the ML estimates of λ and d , in finite samples. With this aim, we simulated 1000 series of length $n=500$ (a common size in the research area where long memory processes are applied) from ARFIMA(0, d , 0) processes with $d=0.1, 0.3, 0.45$, and $\lambda=0, 0.5, 1$, respectively.

Figure 2 (where the kernel histograms of the corresponding distributions are illustrated) shows that the variance of the ML estimator of λ increases with λ , for any values of d ; instead, the estimator of d has the same variability for any fixed value of λ . Anyway, both the ML estimators have a Normal distribution, with small asymmetry. Finally, as far as concerns the bias of the estimates, our results show that the estimation procedure reveals no sensible effect for both d and λ .

Figure 2: Monte Carlo experiment for the joint estimates of d and λ .



Moreover, since the simulated distribution of $\hat{d} - d$ exhibits a constant variance ranging in $[0.037^2, 0.038^2]$, we conjecture that this could be estimated by: $p^2 \left[6 \sum_{j=1}^m (v_j - \bar{v})^2 \right]^{-1}$,

where $v_j = \log(1 - \cos(\mathbf{w}_j))$, $\bar{v} = \frac{1}{m} \sum_{j=1}^m v_j$, as it happens for the least squares estimator of d computed on the spectral ordinates (Brockwell and Davis, 1991, p. 530). We verified that the conjectured expression satisfies the observed range for the variance when we let $m = n^{0.84}$. Further work will deepen these results in order to derive simple expressions for the asymptotic variances of the proposed estimators.

References

- Beran J. (1994) *Statistics for long memory processes*, Chapman and Hall, New York.
- Brockwell P. J. and Davis R. A. (1991) *Time Series: Theory and Methods*, (2nd edition), Springer-Verlag, New York.
- Brubacher S. R. (1976) *Time series modelling with instantaneous nonlinear transformations*, Ph. D. Thesis, University of Lancaster.
- Corduas M. (2000) Preliminary estimation of ARFIMA models, in: *COMPSTAT 2000*, Physica-Verlag, Heidelberg, 247-252.
- Granger C.W.J. and Joyeaux J. (1980) An introduction to long memory time series models and fractional differencing, *Journal of Time Series Analysis*, 1, 15-30.
- Granger C.W.J. and Newbold P. (1976) Forecasting transformed series, *Journal of the Royal Statistical Society B*, 2, 189-203.
- Hosking J.R.M. (1981) Fractional differencing, *Biometrika*, 68, 165-176.
- Xia Y., Tong H., Li W.K., Zhu L. (2000) On the estimation of an instantaneous transformation for time series, *Journal of the Royal Statistical Society B*, 62, 383-397.