Computational Statistics and Data Analysis 53 (2009) 2089–2106

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

journal homepage: www.elsevier.com/locate/csda

Type I and type II fractional Brownian motions: A reconsideration

James Davidson [a,*], Nigar Hashimzade [b]

[a] Department of Economics, University of Exeter, Streatham Court, Rennes Drive, Exeter EX4 4PU, UK
[b] University of Reading, UK

ARTICLE INFO

Article history:
Available online 19 November 2008

ABSTRACT

The so-called type I and type II fractional Brownian motions are limit distributions associated with the fractional integration model in which pre-sample shocks are either included in the lag structure, or suppressed. There can be substantial differences between the distributions of these two processes and of functionals derived from them, so that it becomes an important issue to decide which model to use as a basis for inference. Alternative methods for simulating the type I case are contrasted, and for models close to the nonstationarity boundary, truncating infinite sums is shown to result in a significant distortion of the distribution. A simple simulation method that overcomes this problem is described and implemented. The approach also has implications for the estimation of type I ARFIMA models, and a new conditional ML estimator is proposed, using the annual Nile minima series for illustration.

1. Introduction

The literature on long memory processes in econometrics (for recent examples, see *inter alia* Johansen and Nielsen (2008), Caporale and Gil-Alana (2008), Coakley et al. (2008) and Haldrup and Nielsen (2007)) has adopted two distinct models as a basis for the asymptotic analysis, the limit processes specified being known respectively as type I and type II fractional Brownian motion (fBM). These processes have been carefully examined and contrasted by Marinucci and Robinson (1999). When considered as real continuous processes on the unit interval, they can be defined respectively by

$$X(r) = \frac{1}{\Gamma(d+1)}\int_0^r (r-s)^d \mathrm{d}B(s) + \frac{1}{\Gamma(d+1)}\int_{-\infty}^0 [(r-s)^d - (-s)^d]\mathrm{d}B(s) \tag{1.1}$$

and

$$X^*(r) = \frac{1}{\Gamma(d+1)}\int_0^r (r-s)^d \mathrm{d}B(s), \tag{1.2}$$

where $-\frac{1}{2} < d < \frac{1}{2}$ and B denotes regular Brownian motion. In other words, in the type II case the second term in (1.1) is omitted. It will be convenient to write the decomposition

$$X = X^* + X^{**}, \tag{1.3}$$

where $X^{**}(r)$ is defined as the second of the two terms in (1.1). The processes X^* and X^{**} are Gaussian, and independent of each other, so we know that the variance of (1.1) will exceed that of (1.2). As shown by Marinucci and Robinson (1999), the increments of (1.1) are stationary, whereas those of (1.2) are not.

* Corresponding author. Tel.: +44 01392 264517; fax: +44 07092 812113.
E-mail address: james.davidson@exeter.ac.uk (J. Davidson).

0167-9473/$ – see front matter
doi:10.1016/j.csda.2008.11.008

These processes are commonly motivated by postulating realizations of size n of discrete processes and considering the weak limits of normalized partial sums, as $n \to \infty$. Define

$$x_t = (1 - L)^{-d} u_t, \tag{1.4}$$

where we assume for the sake of exposition that $\{u_t\}_{-\infty}^{\infty}$ is an i.i.d. process with mean 0 and variance σ^2, and

$$(1 - L)^{-d} = \sum_{j=0}^{\infty} b_j L^j, \tag{1.5}$$

where, letting $\Gamma(\cdot)$ denote the gamma function,

$$b_j = \frac{\Gamma(d + j)}{\Gamma(d)\Gamma(1 + j)}. \tag{1.6}$$

Defining the partial sum process

$$X_n(r) = \frac{1}{\sigma n^{1/2+d}} \sum_{t=1}^{[nr]} x_t \tag{1.7}$$

it is known that $X_n \xrightarrow{d} X$, where $\xrightarrow{d}$ denotes weak convergence in the space of measures on $D_{[0,1]}$, the space of cadlag functions of the unit interval equipped with the Skorokhod topology. (See for example Davidson and de Jong (2000)). On the other hand, defining

$$u_t^* = 1(t > 0)u_t \tag{1.8}$$

and x_t^* as the case corresponding to x_t in (1.4) when u_t^* replaces u_t, and then defining X_n^* like (1.7) with x_t^* replacing x_t, it is known that $X_n^* \xrightarrow{d} X^*$ (Marinucci and Robinson, 2000).

The model in (1.8) is one that is often used in simulation exercises to generate fractionally integrated processes, as an alternative to the procedure of setting a fixed, finite truncation of the lag distribution in (1.4), common to every t. However, from the point of view of modelling real economic or financial time series, model (1.8) is obviously problematic. There is, in most cases, nothing about the date when we start to observe a series which suggests that we ought to set all shocks preceding it to 0. Such truncation procedures are common in time series modelling, but are usually justified by the assumption that the effect is asymptotically negligible. In this case, however, where the effect is manifestly not negligible in the limit, the choice of model becomes a critical issue.

The setting for this choice is the case where a Monte Carlo simulation is to be used to construct the null distribution of a test statistic postulated to be a functional of fBM. If model (1.8) is used to generate the artificial data, then the distribution so simulated will be the Type II case. However, if the observed data ought to be treated as drawn from (1.4), then the estimated critical values will be incorrect even in large samples. It then becomes of importance to know how large this error is.

Section 2 of the paper reviews and contrasts the main properties of these models. A leading difficulty in working with the type I model is to simulate it effectively, and as we show in Section 3, the fixed lag truncation strategy is not generally effective, except by expending a dramatically large amount of computing resources. Since type I fBM has a harmonizable representation, another suggestion has been to use this to simulate the model, and then use a fast Fourier transform to recover the data in the time domain. However, we also show that this method cannot function effectively without large resources. Methods for generating type I processes do exist, for example circulant embedding and wavelet aproximations, but these are relatively difficult to implement in an econometric context. In Section 4 we suggest a new simulation method for type I processes, whose computational demands are trivial, and being implemented in the time domain adapts naturally to econometric modelling applications. The method is highly accurate when the data are Gaussian, and is always asymptotically valid.

Finally, we point out in Section 5 how the same approximation technique can be used to estimate ARFIMA time series models under the assumption that the true processes are of type I. This is in contrast to the usual time domain estimation by least squares, or conditional maximum likelihood, where the necessity of truncating lag distributions to match the observed data series implicitly (and perhaps inappropriately) imposes restrictions appropriate to the type II case. The method entails fitting some constructed regressors, whose omission will potentially bias the estimates in finite samples. The technique is illustrated with an application to the well-known series of annual Nile minima. Section 6 concludes the paper. Proofs are contained in Appendix A, and Appendix B exhibits some simulations of representative fractional Brownian functionals, under the two definitions.

The computations in this paper were carried out using the package Time Series Modelling 4 (Davidson, 2008) which runs under the Ox 4/5 matrix programming system (Doornik, 2006).