Estimation of Possibly Non-Stationary Long Memory Processes via Adaptive Overdifferencing

Maryclare Griffin

University of Massachusetts Amherst
Department of Mathematics and Statistics

May 20, 2023

Joint work with David S. Matteson, Gennady Samorodnitsky (Cornell)

Hello and congratulations, Ruey!

Journal of Forecasting, Vol. 13, 109–131 (1994)

Biometrika (1997), **84**, 4, pp. 791–802
Printed in Great Britain

Some Advances in Non-linear and Adaptive Modelling in Time-series

GEORGE C. TIAO AND RUEY S. TSAY
University of Chicago, IL, U.S.A.

Bandwidth selection for kernel regression with long-range dependent errors

By BONNIE K. RAY
Center for Applied Mathemat
nology, Newark, New Jersey (
e-mail: borayx@m.njit.edu
AND RUEY S. TSAY

True or Spurious Long Memory? A New Test

Arek Ohanissian, Jeffrey R. Russell, and Ruey S. Tsay
Graduate School of Business, University of Chicago, Chicago, IL 60637 (*jeffrey.russell@chicagogsb.edu*)

It is well known that long memory characteristics observed in data can be generated by nonstationary structural-break or slow regime switching models. We propose a statistical test to distinguish between true long memory and spurious long memory based on invariance of the long memory parameter for temporal aggregates of the process under the null of true long memory. Geweke Porter-Hudak estimates of the long memory parameter obtained from different temporal aggregates of the underlying time series are shown to be asymptotically jointly normal, leading to a test statistic that is constructed as the quadratic form of a demeaned vector of the estimates. The result is a test statistic that is very simple to implement. Simulations show the test to have good size and power properties for the classic alternatives to true long memory that have been suggested in the literature. The asymptotic distribution of the test statistic is also valid for a stochastic volatility with Gaussian long memory model. The test is applied to foreign exchange rate data. Based on all the models considered in this article, we conclude that the long memory property in exchange rate volatility is generated by a true long memory process.

KEY WORDS: Regime switching; Structural change; Temporal aggregation.

BAYESIAN METHODS FOR CHANGE-POINT DETECTION IN LONG-RANGE DEPENDENT PROCESSES

By Bonnie K. Ray and Ruey S. Tsay

Iterative Bandwidth Estimation for Nonparametric Regression with Long-range Dependent Errors

Bonnie K. Ray
Ruey S. Tsay

ABSTRACT We discuss the problem of bandwidth selection for a kernel
or when the errors are long-range dependent. The
dth selection method is investigated and modified
ory in the errors. We compare the mean average-
nd estimates using the bandwidth obtained from
to that obtained assuming short-range dependent
ve apply the modified method to estimate a trend

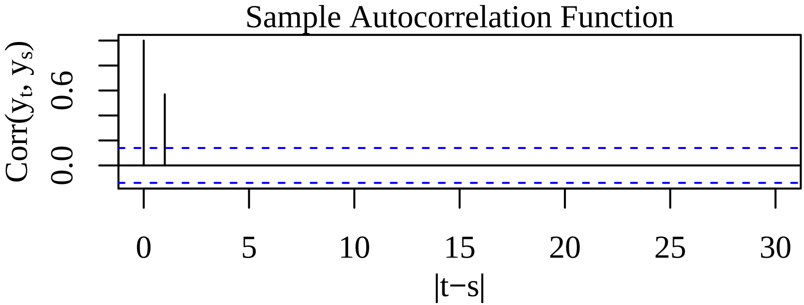
Long-range Dependence in Daily Stock Volatilities

Bonnie K. Ray
Department of Mathematical Sciences and Center for Applied Math and Statistics, New Jersey Institute of Technology, Newark, NJ 07102 (*borayx@m.njit.edu*)

Ruey S. Tsay
Graduate School of Business, University of Chicago, Chicago, IL 60637 (*ruey.tsay@gsb.uchicago.edu*)

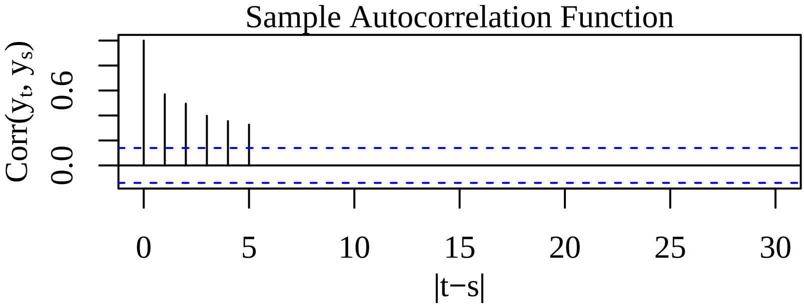
Long Memory

Corr (y_t, y_s) decreases very slowly as $|t-s|$ increases



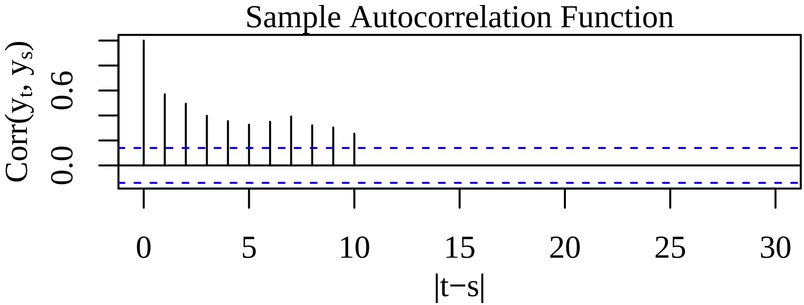
Long Memory

$\text{Corr}\,(y_t, y_s)$ decreases very slowly as $|t-s|$ increases



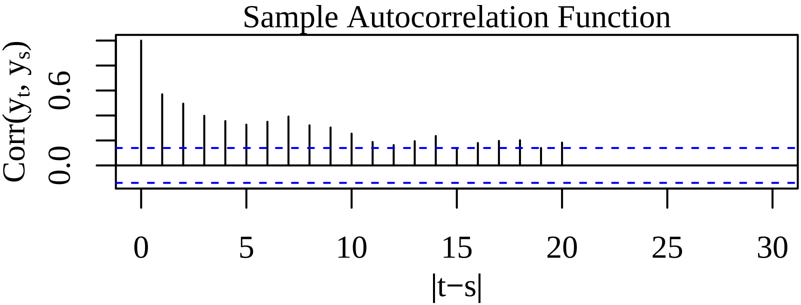
Long Memory

Corr (y_t, y_s) decreases very slowly as $|t - s|$ increases



Long Memory

$\mathrm{Corr}(y_t, y_s)$ decreases very slowly as $|t - s|$ increases



AR and Long Memory

AR and Long Memory

$$(y_t - \mu_t) = \sum_{j=1}^{p} \phi_j \left(y_{t-j} - \mu_{t-j}\right) + z_t$$

$$z_t \overset{i.i.d.}{\sim} \text{normal}\left(0, \sigma^2\right)$$

AR and Long Memory

$$(y_t - \mu_t) = \sum_{j=1}^{p} \phi_j \left(y_{t-j} - \mu_{t-j}\right) + z_t$$

$$z_t \overset{i.i.d.}{\sim} \text{normal}\left(0, \sigma^2\right)$$

Long memory $\implies$ $|\phi_j|$ not "small" for large p, effectively

AR and Long Memory

$$(y_t - \mu_t) = \sum_{j=1}^{p} \phi_j \left(y_{t-j} - \mu_{t-j}\right) + z_t$$

$$z_t \overset{i.i.d.}{\sim} \text{normal}\left(0, \sigma^2\right)$$

Long memory $\implies$ $|\phi_j|$ not "small" for large p, effectively

$$(y_t - \mu_t) = \sum_{j=1}^{\infty} \phi_j \left(y_{t-j} - \mu_{t-j}\right) + z_t$$

$$z_t \overset{i.i.d.}{\sim} \text{normal}\left(0, \sigma^2\right)$$

AR and Long Memory

$$(y_t - \mu_t) = \sum_{j=1}^{p} \phi_j \left(y_{t-j} - \mu_{t-j}\right) + z_t$$

$$z_t \overset{i.i.d.}{\sim} \text{normal}\left(0, \sigma^2\right)$$

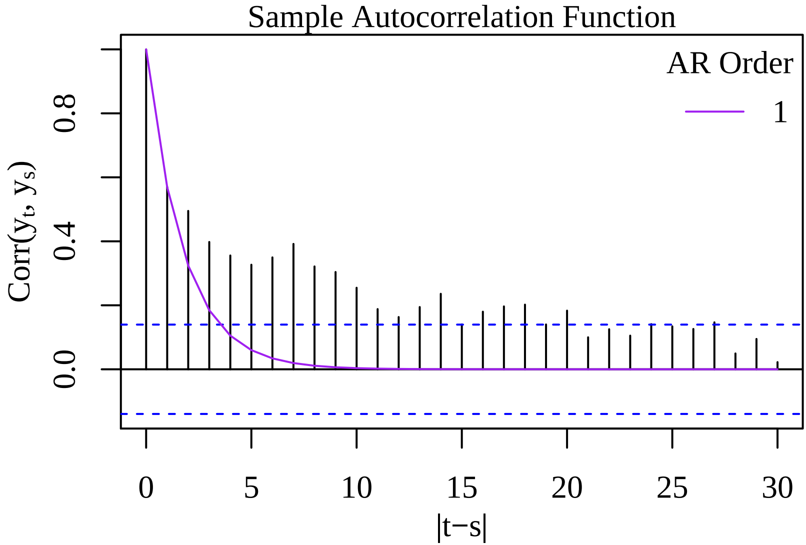
Long memory $\implies$ $|\phi_j|$ not "small" for large p, effectively

$$(y_t - \mu_t) = \sum_{j=1}^{\infty} \phi_j \left(y_{t-j} - \mu_{t-j}\right) + z_t$$

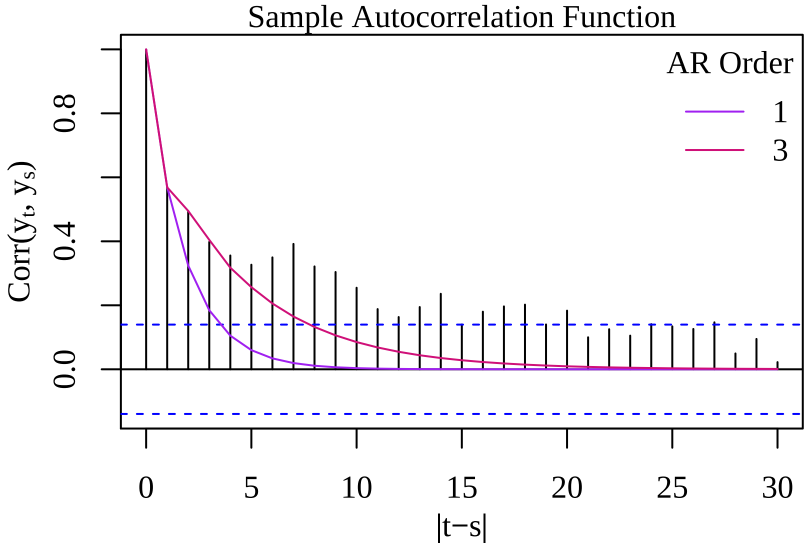
$$z_t \overset{i.i.d.}{\sim} \text{normal}\left(0, \sigma^2\right)$$

In practice, long memory requires $p \to \infty$ as $n \to \infty$

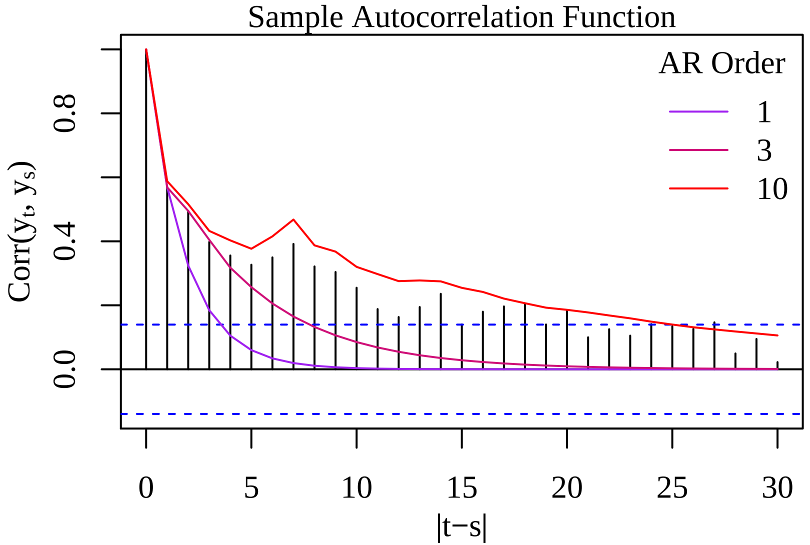
AR and Long Memory



AR and Long Memory



AR and Long Memory



Fractional Integration (FI)

$$\overbrace{(y_t - \mu_t) = \sum_{j=1}^{\infty} \binom{d}{j} (-1)^{j-1} \left(y_{t-j} - \mu_{t-j}\right) + z_t}^{\text{FI}(d)}$$

$$z_t \overset{i.i.d.}{\sim} \text{normal}\left(0, \sigma^2\right)$$

Granger and Joyeux (1980); Hosking (1981)

Fractional Integration (FI)

$$\overbrace{(y_t - \mu_t) = \sum_{j=1}^{\infty} \binom{d}{j} (-1)^{j-1} (y_{t-j} - \mu_{t-j}) + z_t}^{\text{FI}(d)}$$

$$z_t \overset{i.i.d.}{\sim} \text{normal}\left(0, \sigma^2\right)$$

Can be written $(1 - B)^d (y_t - \mu_t) = z_t$

Granger and Joyeux (1980); Hosking (1981)

Fractional Integration (FI)

$$\overbrace{(y_t - \mu_t) = \sum_{j=1}^{\infty} \binom{d}{j} (-1)^{j-1} (y_{t-j} - \mu_{t-j}) + z_t}^{\text{FI}(d)}$$

$$z_t \overset{i.i.d.}{\sim} \text{normal}\left(0, \sigma^2\right)$$

Can be written $(1 - B)^d (y_t - \mu_t) = z_t$

Stationary if $d < 0.5$

Granger and Joyeux (1980); Hosking (1981)

Fractional Integration (FI)

$$\overbrace{(y_t - \mu_t) = \sum_{j=1}^{\infty} \binom{d}{j} (-1)^{j-1} (y_{t-j} - \mu_{t-j}) + z_t}^{\mathrm{FI}(d)}$$

$$z_t \overset{i.i.d.}{\sim} \text{normal}\left(0, \sigma^2\right)$$

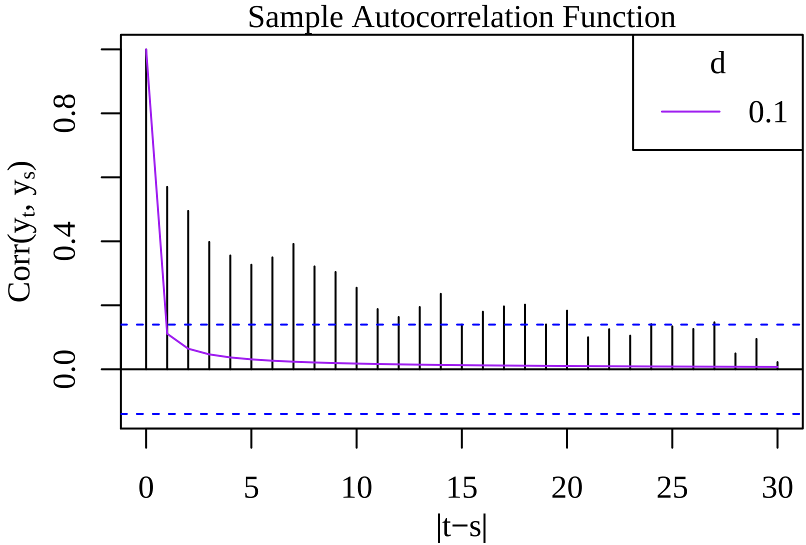
Can be written $(1 - B)^d (y_t - \mu_t) = z_t$

Stationary if $d < 0.5$

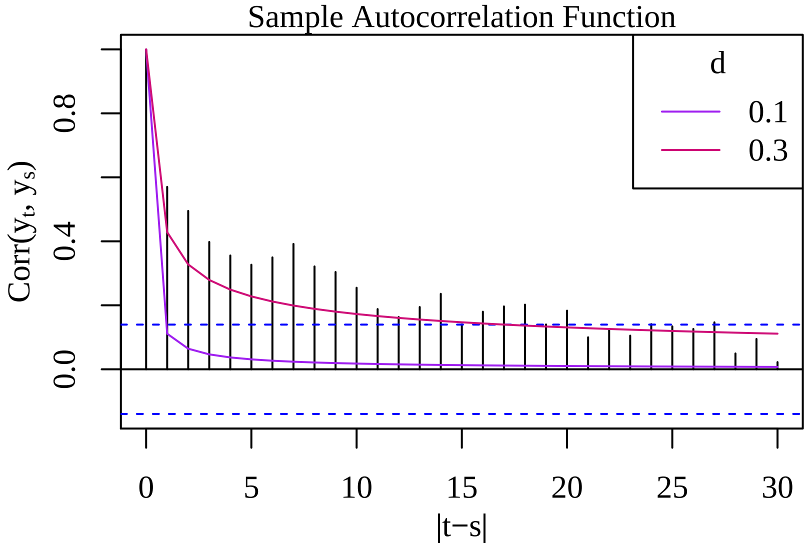
Unique values of $d \le -0.5$ give unique likelihood

Granger and Joyeux (1980); Hosking (1981)

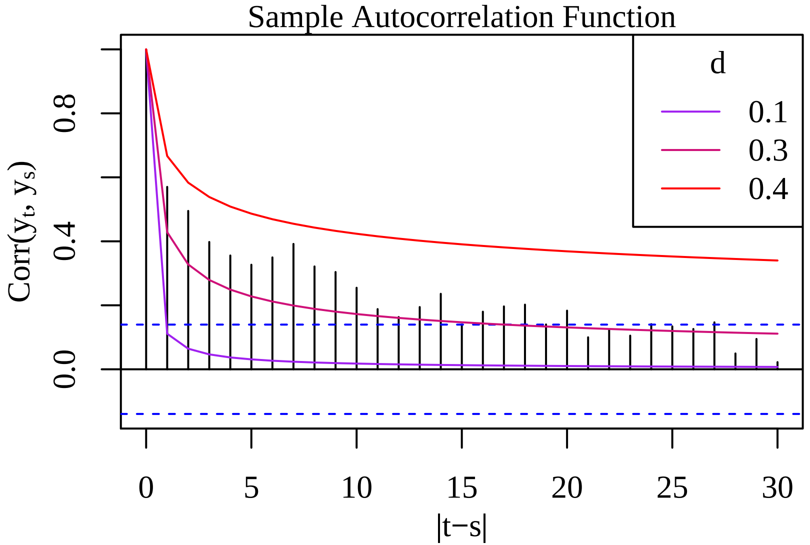
Fractional Integration (FI)



Fractional Integration (FI)



Fractional Integration (FI)



Autoregressive FI Moving Average (ARFIMA) Model

$$\overbrace{\phi\left(B\right)\left(1-B\right)^{d}\left(y_t-\mu_t\right)=\theta\left(B\right)z_t, \quad z_t \overset{i.i.d.}{\sim} \text{normal}\left(0,\sigma^2\right)}^{\text{ARFIMA}(p,d,q)}$$

Autoregressive FI Moving Average (ARFIMA) Model

$$\overbrace{\phi\left(B\right)\left(1-B\right)^{d}\left(y_t-\mu_t\right)=\theta\left(B\right)z_t, \quad z_t \overset{i.i.d.}{\sim} \text{normal}\left(0,\sigma^2\right)}^{\text{ARFIMA}(p,d,q)}$$

Constrained to be

Autoregressive FI Moving Average (ARFIMA) Model

$$\overbrace{\phi\left(B\right)\left(1-B\right)^{d}\left(y_t-\mu_t\right)=\theta\left(B\right)z_t, \quad z_t \overset{i.i.d.}{\sim} \text{normal}\left(0,\sigma^2\right)}^{\text{ARFIMA}(p,d,q)}$$

Constrained to be

- **Stationary:** Roots of $\phi\left(x\right)$ outside unit circle, $d<0.5$

Autoregressive FI Moving Average (ARFIMA) Model

$$\overbrace{\phi(B)(1-B)^d(y_t-\mu_t)=\theta(B)\,z_t, \quad z_t \overset{i.i.d.}{\sim} \text{normal}\left(0,\sigma^2\right)}^{\text{ARFIMA}(p,d,q)}$$

Constrained to be

- **Stationary:** Roots of $\phi(x)$ outside unit circle, $d < 0.5$
- **Invertible:** Roots of $\theta(x)$ outside unit circle, $d \geq -0.5$

Autoregressive FI Moving Average (ARFIMA) Model

$$\overbrace{\phi(B)(1-B)^d(y_t-\mu_t)=\theta(B)z_t, \quad z_t \overset{i.i.d.}{\sim} \text{normal}\left(0,\sigma^2\right)}^{\text{ARFIMA}(p,d,q)}$$

Constrained to be

- **Stationary:** Roots of $\phi(x)$ outside unit circle, $d < 0.5$
- **Invertible:** Roots of $\theta(x)$ outside unit circle, $d \geq -0.5$

This talk: Focus on FI(d) =ARFIMA$(0, d, 0)$

Autoregressive FI Moving Average (ARFIMA) Model

$$\overbrace{\phi\left(B\right)\left(1-B\right)^{d}\left(y_t-\mu_t\right)=\theta\left(B\right)z_t, \quad z_t \overset{i.i.d.}{\sim} \text{normal}\left(0,\sigma^2\right)}^{\text{ARFIMA}(p,d,q)}$$

Constrained to be

- **Stationary:** Roots of $\phi\left(x\right)$ outside unit circle, $d<0.5$
- **Invertible:** Roots of $\theta\left(x\right)$ outside unit circle, $d \geq -0.5$

This talk: Focus on FI(d) =ARFIMA$(0,d,0)$

(Content applies to ARFIMA(p,d,q) as well)

Estimation of ARFIMA(p, d, q) Processes

Baillie (1996)

Estimation of ARFIMA(p, d, q) Processes

Lots of methods, based on:

Baillie (1996)

Estimation of ARFIMA(p, d, q) Processes

Lots of methods, based on:

- Regression

Baillie (1996)

Estimation of ARFIMA(p, d, q) Processes

Lots of methods, based on:

- Regression
  - ☺ Doesn't have to make assumptions on range of d

Baillie (1996)

Estimation of ARFIMA(p, d, q) Processes

Lots of methods, based on:

- Regression
  - ☺ Doesn't have to make assumptions on range of d
  - ☹ "Throws out" data, may not give estimates of all parameters

Baillie (1996)

Estimation of ARFIMA(p, d, q) Processes

Lots of methods, based on:

- Regression
  - ☺ Doesn't have to make assumptions on range of d
  - ☹ "Throws out" data, may not give estimates of all parameters
  - ☹ Often introduces additional tuning parameters

Baillie (1996)

Estimation of ARFIMA(p, d, q) Processes

Lots of methods, based on:

- Regression
  - ☺ Doesn't have to make assumptions on range of d
  - ☹ "Throws out" data, may not give estimates of all parameters
  - ☹ Often introduces additional tuning parameters
  - ☹ Difficult to get standard errors, confidence intervals for d

Baillie (1996)

Estimation of ARFIMA(p, d, q) Processes

Lots of methods, based on:

- Regression
  - ☺ Doesn't have to make assumptions on range of d
  - ☹ "Throws out" data, may not give estimates of all parameters
  - ☹ Often introduces additional tuning parameters
  - ☹ Difficult to get standard errors, confidence intervals for d
- Likelihood

Baillie (1996)

Estimation of ARFIMA(p, d, q) Processes

Lots of methods, based on:

- Regression
  - ☺ Doesn't have to make assumptions on range of d
  - ☹ "Throws out" data, may not give estimates of all parameters
  - ☹ Often introduces additional tuning parameters
  - ☹ Difficult to get standard errors, confidence intervals for d
- Likelihood
  - ☹ Requires assumptions on range of d

Baillie (1996)

Estimation of ARFIMA(p, d, q) Processes

Lots of methods, based on:

- Regression
  - ☺ Doesn't have to make assumptions on range of d
  - ☹ "Throws out" data, may not give estimates of all parameters
  - ☹ Often introduces additional tuning parameters
  - ☹ Difficult to get standard errors, confidence intervals for d
- Likelihood
  - ☹ Requires assumptions on range of d
  - ☹ Needs more computational tricks

Baillie (1996)

Estimation of ARFIMA(p, d, q) Processes

Lots of methods, based on:

- Regression
  - ☺ Doesn't have to make assumptions on range of d
  - ☹ "Throws out" data, may not give estimates of all parameters
  - ☹ Often introduces additional tuning parameters
  - ☹ Difficult to get standard errors, confidence intervals for d
- Likelihood
  - ☹ Requires assumptions on range of d
  - ☹ Needs more computational tricks
  - ☺ Theory for standard errors, confidence intervals applies

Baillie (1996)

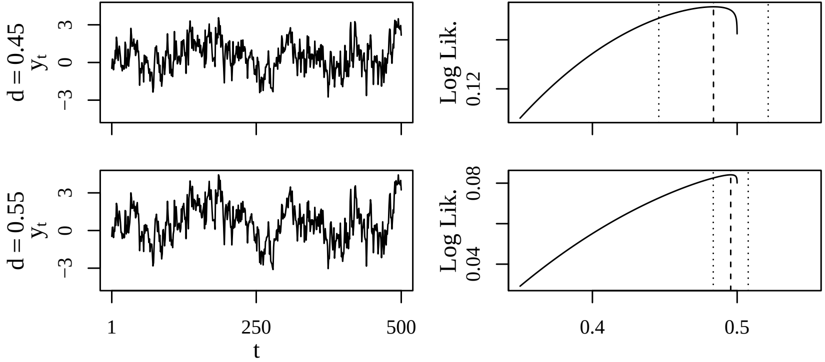
Estimation of ARFIMA(p, d, q) Processes

Lots of methods, based on:

- Regression
  - ☺ Doesn't have to make assumptions on range of d
  - ☹ "Throws out" data, may not give estimates of all parameters
  - ☹ Often introduces additional tuning parameters
  - ☹ Difficult to get standard errors, confidence intervals for d
- Likelihood
  - ☹ **Requires assumptions on range of d**
  - ☹ Needs more computational tricks
  - ☺ Theory for standard errors, confidence intervals applies

Baillie (1996)

Why do we care about assumptions on d?



Can we get around this?

Can we get around this?

Yes - because polynomials in B are just like regular polynomials

Can we get around this?

Yes - because polynomials in B are just like regular polynomials

$$(1 - B)^{d-m} (1 - B)^{m} (y_t - \mu_t) = z_t$$

Can we get around this?

Yes - because polynomials in B are just like regular polynomials

$$(1 - B)^{d-m} (1 - B)^m (y_t - \mu_t) = z_t$$

- If $0.5 \leq d < 1.5$, choose $m = 1$ to get a stationary process

Can we get around this?

Yes - because polynomials in B are just like regular polynomials

$$(1 - B)^{d-m} (1 - B)^m (y_t - \mu_t) = z_t$$

- If $0.5 \leq d < 1.5$, choose $m = 1$ to get a stationary process

$$(1 - B) (y_t - \mu_t) = y_t - y_{t-1} - (\mu_t - \mu_{t-1})$$

Can we get around this?

Yes - because polynomials in B are just like regular polynomials

$$(1-B)^{d-m}(1-B)^{m}(y_t-\mu_t)=z_t$$

- If $0.5 \leq d < 1.5$, choose $m = 1$ to get a stationary process

$$(1-B)(y_t-\mu_t)=y_t-y_{t-1}-(\mu_t-\mu_{t-1})$$

- If $1.5 \leq d < 2.5$, choose $m = 2$ to get a stationary process

Can we get around this?

Yes - because polynomials in B are just like regular polynomials

$$(1-B)^{d-m}(1-B)^{m}(y_t-\mu_t)=z_t$$

- If $0.5 \leq d < 1.5$, choose $m=1$ to get a stationary process

$$(1-B)(y_t-\mu_t)=y_t-y_{t-1}-(\mu_t-\mu_{t-1})$$

- If $1.5 \leq d < 2.5$, choose $m=2$ to get a stationary process

$$(1-B)^{2}(y_t-\mu_t)=y_t-2y_{t-1}+y_{t-2}-(\mu_t-2\mu_{t-1}+\mu_{t-2})$$

Can we get around this?

Suggests the following procedure

Can we get around this?

Suggests the following procedure

1. For $d < 0.5$, fit an $\text{FI}(d)$ model to y_t

Can we get around this?

Suggests the following procedure

1. For $d < 0.5$, fit an $\mathrm{FI}(d)$ model to y_t
2. For $0.5 \leq d < 1.5$, fit an $\mathrm{FI}(d-1)$ model to $(1-B)\, y_t$

Can we get around this?

Suggests the following procedure

1. For $d < 0.5$, fit an $\mathrm{FI}(d)$ model to y_t
2. For $0.5 \leq d < 1.5$, fit an $\mathrm{FI}(d-1)$ model to $(1-B)\,y_t$
3. For $1.5 \leq d < 2.5$, fit an $\mathrm{FI}(d-2)$ model to $(1-B)^2\,y_t$

Can we get around this?

Suggests the following procedure

1. For $d < 0.5$, fit an $\mathrm{FI}(d)$ model to y_t
2. For $0.5 \leq d < 1.5$, fit an $\mathrm{FI}(d-1)$ model to $(1-B)\, y_t$
3. For $1.5 \leq d < 2.5$, fit an $\mathrm{FI}(d-2)$ model to $(1-B)^2\, y_t$

Select $d \leq 2.5$ that maximizes likelihood

Can we get around this?

Suggests the following procedure

1. For $d < 0.5$, fit an FI(d) model to y_t
2. For $0.5 \leq d < 1.5$, fit an FI($d - 1$) model to $(1 - B)\, y_t$
3. For $1.5 \leq d < 2.5$, fit an FI($d - 2$) model to $(1 - B)^2\, y_t$

Select $d \leq 2.5$ that maximizes likelihood

Problem: Likelihoods in each step based on different data!

Can we get around this?

Suggests the following procedure

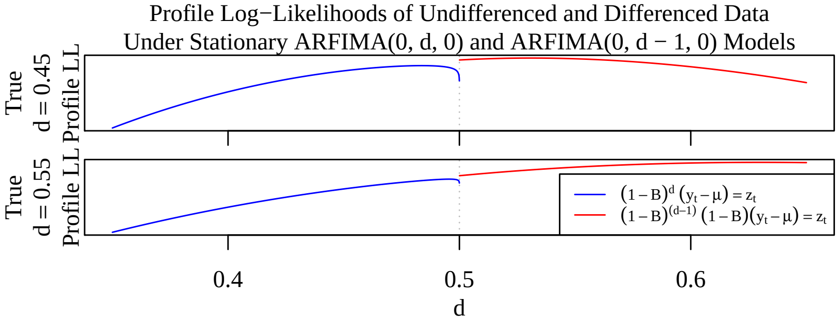
1. For $d < 0.5$, fit an $\mathrm{FI}(d)$ model to y_t
2. For $0.5 \leq d < 1.5$, fit an $\mathrm{FI}(d-1)$ model to $(1-B)\, y_t$
3. For $1.5 \leq d < 2.5$, fit an $\mathrm{FI}(d-2)$ model to $(1-B)^2\, y_t$

Select $d \leq 2.5$ that maximizes likelihood

Problem: Likelihoods in each step based on different data!

Beran (1995); Velasco and Robinson (2000); Mayoral (2007)

What's wrong with this procedure?



Is there anything else we can do?

Is there anything else we can do?

- **Problem:** Data changes as d changes

Is there anything else we can do?

- **Problem:** Data changes as d changes
- **Solution:** Fix data in advance to get stationarity for $d < \bar{d}$

Key Idea

Odaki (1993); Doornik and Ooms (2003)

Key Idea

- Given $\bar{d}$, we can find m such that $\bar{d} - m < 0.5$

Odaki (1993); Doornik and Ooms (2003)

Key Idea

- Given $\bar{d}$, we can find m such that $\bar{d} - m < 0.5$
- As long as $d < \bar{d}$, $(1 - B)^m \, y_t$ is stationary FI$(d - m)$

Odaki (1993); Doornik and Ooms (2003)

Key Idea

- Given $\bar{d}$, we can find m such that $\bar{d} - m < 0.5$
- As long as $d < \bar{d}$, $(1 - B)^m \, y_t$ is stationary FI$(d - m)$

☹ Most software won't give likelihood if $d - m \leq -0.5$

Odaki (1993); Doornik and Ooms (2003)

Key Idea

- Given $\bar{d}$, we can find m such that $\bar{d} - m < 0.5$
- As long as $d < \bar{d}$, $(1-B)^m\, y_t$ is stationary FI$(d-m)$

☹ Most software won't give likelihood if $d - m \leq -0.5$

$$(1-B)^{d-m}\,(1-B)^m\,(y_t - \mu_t) = z_t,\; z_t \overset{i.i.d.}{\sim} \text{normal}\left(0, \sigma^2\right)$$

Odaki (1993); Doornik and Ooms (2003)

Key Idea

- Given $\bar{d}$, we can find m such that $\bar{d} - m < 0.5$
- As long as $d < \bar{d}$, $(1-B)^m\, y_t$ is stationary FI$(d-m)$
- ☹ Most software won't give likelihood if $d - m \leq -0.5$

$$(1-B)^{d-m}\,(1-B)^m\,(y_t - \mu_t) = z_t,\; z_t \overset{i.i.d.}{\sim} \text{normal}\left(0, \sigma^2\right)$$

- ☺ Can find integer r such that $d - m + r > -0.5$

Odaki (1993); Doornik and Ooms (2003)

Key Idea

- Given $\bar{d}$, we can find m such that $\bar{d} - m < 0.5$
- As long as $d < \bar{d}$, $(1-B)^m y_t$ is stationary FI$(d-m)$

☹ Most software won't give likelihood if $d - m \leq -0.5$

$$(1-B)^{d-m}(1-B)^m(y_t - \mu_t) = z_t,\ z_t \overset{i.i.d.}{\sim} \text{normal}\left(0, \sigma^2\right)$$

☺ Can find integer r such that $d - m + r > -0.5$

$$(1-B)^{d-m+r}(1-B)^m(y_t - \mu_t) = (1-B)^r z_t$$

Odaki (1993); Doornik and Ooms (2003)

Key Idea

- Given $\bar{d}$, we can find m such that $\bar{d} - m < 0.5$
- As long as $d < \bar{d}$, $(1-B)^m y_t$ is stationary FI$(d-m)$

☹ Most software won't give likelihood if $d - m \leq -0.5$

$$(1-B)^{d-m}(1-B)^m(y_t - \mu_t) = z_t,\ z_t \overset{i.i.d.}{\sim} \text{normal}\left(0, \sigma^2\right)$$

☺ Can find integer r such that $d - m + r > -0.5$

$$(1-B)^{d-m+r}(1-B)^m(y_t - \mu_t) = (1-B)^r z_t$$

$(1-B)^m(y_t - \mu_t)$ stationary, constrained ARFIMA$(0, d-m+r, r)$

Odaki (1993); Doornik and Ooms (2003)

New Procedure

New Procedure

- Pick $\bar{d}$, maximum possible d

New Procedure

- Pick $\bar{d}$, maximum possible d
- Find integer m such that $-0.5 \leq \bar{d} - m \leq 0.5$

New Procedure

- Pick $\bar{d}$, maximum possible d
- Find integer m such that $-0.5 \leq \bar{d} - m \leq 0.5$
- Compute $w_t = (1 - B)^m \, y_t$

New Procedure

- Pick $\bar{d}$, maximum possible d
- Find integer m such that $-0.5 \leq \bar{d} - m \leq 0.5$
- Compute $w_t = (1 - B)^m \, y_t$
- Evaluate sequence of stationary ARFIMA likelihoods

New Procedure

- Pick $\bar{d}$, maximum possible d
- Find integer m such that $-0.5 \leq \bar{d} - m \leq 0.5$
- Compute $w_t = (1 - B)^m \, y_t$
- Evaluate sequence of stationary ARFIMA likelihoods
    1. For $-0.5 \leq d - m < 0.5$, $w_t \sim$ARFIMA$(0, d - m, 0)$

New Procedure

- Pick $\bar{d}$, maximum possible d
- Find integer m such that $-0.5 \leq \bar{d} - m \leq 0.5$
- Compute $w_t = (1 - B)^m \, y_t$
- Evaluate sequence of stationary ARFIMA likelihoods
    1. For $-0.5 \leq d - m < 0.5$, $w_t \sim$ARFIMA$(0, d - m, 0)$
    2. For $-1.5 \leq d - m < -0.5$, $w_t \sim$ARFIMA$(0, d - m + 1, 1)$

New Procedure

- Pick $\bar{d}$, maximum possible d
- Find integer m such that $-0.5 \leq \bar{d} - m \leq 0.5$
- Compute $w_t = (1 - B)^m \, y_t$
- Evaluate sequence of stationary ARFIMA likelihoods
  1. For $-0.5 \leq d - m < 0.5$, $w_t \sim$ARFIMA$(0, d - m, 0)$
  2. For $-1.5 \leq d - m < -0.5$, $w_t \sim$ARFIMA$(0, d - m + 1, 1)$
  3. For $-2.5 \leq d - m < -1.5$, $w_t \sim$ARFIMA$(0, d - m + 2, 2)$

New Procedure

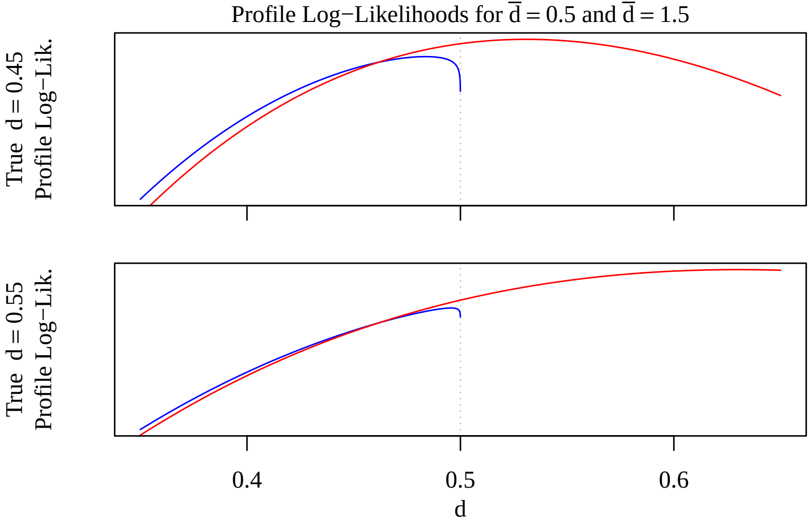
- Pick $\bar{d}$, maximum possible d
- Find integer m such that $-0.5 \leq \bar{d} - m \leq 0.5$
- Compute $w_t = (1 - B)^m \, y_t$
- Evaluate sequence of stationary ARFIMA likelihoods
    1. For $-0.5 \leq d - m < 0.5$, $w_t \sim$ARFIMA$(0, d - m, 0)$
    2. For $-1.5 \leq d - m < -0.5$, $w_t \sim$ARFIMA$(0, d - m + 1, 1)$
    3. For $-2.5 \leq d - m < -1.5$, $w_t \sim$ARFIMA$(0, d - m + 2, 2)$
- Select $-2.5 + m \leq d \leq 0.5 + m$ that maximizes likelihood

New Procedure

- Pick $\bar{d}$, maximum possible d
- Find integer m such that $-0.5 \leq \bar{d} - m \leq 0.5$
- Compute $w_t = (1 - B)^m y_t$
- Evaluate sequence of stationary ARFIMA likelihoods
  1. For $-0.5 \leq d - m < 0.5$, $w_t \sim$ARFIMA$(0, d - m, 0)$
  2. For $-1.5 \leq d - m < -0.5$, $w_t \sim$ARFIMA$(0, d - m + 1, 1)$
  3. For $-2.5 \leq d - m < -1.5$, $w_t \sim$ARFIMA$(0, d - m + 2, 2)$
- Select $-2.5 + m \leq d \leq 0.5 + m$ that maximizes likelihood

This gives a continuous likelihood in d!

New Procedure



What's left?!

What's left?!

Need a principled/optimal way to choose $\bar{d}$

What's left?!

Need a principled/optimal way to choose $\bar{d}$

Choose smallest $\bar{d}$ for which:

What's left?!

Need a principled/optimal way to choose $\bar{d}$

Choose smallest $\bar{d}$ for which:

- Likelihood is decreasing in d near boundary $\bar{d}$

What's left?!

Need a principled/optimal way to choose $\bar{d}$

Choose smallest $\bar{d}$ for which:

- Likelihood is decreasing in d near boundary $\bar{d}$
- Approximate $(1-\epsilon)\,\%$ quantile of $\hat{d}_{\bar{d}}$ is $< \bar{d}$

What's left?!

Need a principled/optimal way to choose $\bar{d}$

Choose smallest $\bar{d}$ for which:

- Likelihood is decreasing in d near boundary $\bar{d}$
- Approximate $(1-\epsilon)\,\%$ quantile of $\hat{d}_{\bar{d}}$ is $< \bar{d}$

$\Pr(\hat{d}_{\bar{d}} > \bar{d}) \approx \epsilon$, small ϵ facilitates:

What's left?!

Need a principled/optimal way to choose $\bar{d}$

Choose smallest $\bar{d}$ for which:

- Likelihood is decreasing in d near boundary $\bar{d}$
- Approximate $(1-\epsilon)\,\%$ quantile of $\hat{d}_{\bar{d}}$ is $< \bar{d}$

$\Pr(\hat{d}_{\bar{d}} > \bar{d}) \approx \epsilon$, small ϵ facilitates:

- Construction of CI's for d with nominal coverage

What's left?!

Need a principled/optimal way to choose $\bar{d}$

Choose smallest $\bar{d}$ for which:

- Likelihood is decreasing in d near boundary $\bar{d}$
- Approximate $(1-\epsilon)\,\%$ quantile of $\hat{d}_{\bar{d}}$ is $< \bar{d}$

$\Pr(\hat{d}_{\bar{d}} > \bar{d}) \approx \epsilon$, small ϵ facilitates:

- Construction of CI's for d with nominal coverage
- Implementation of tests of $d = c$ with nominal level

What's left?!

Need a principled/optimal way to choose $\bar{d}$

Choose smallest $\bar{d}$ for which:

- Likelihood is decreasing in d near boundary $\bar{d}$
- Approximate $(1-\epsilon)\,\%$ quantile of $\hat{d}_{\bar{d}}$ is $< \bar{d}$

$\Pr(\hat{d}_{\bar{d}} > \bar{d}) \approx \epsilon$, small ϵ facilitates:

- Construction of CI's for d with nominal coverage
- Implementation of tests of $d = c$ with nominal level

When $\epsilon = 0.5$ we'll call estimator **BND**

What's left?!

Need a principled/optimal way to choose $\bar{d}$

Choose smallest $\bar{d}$ for which:

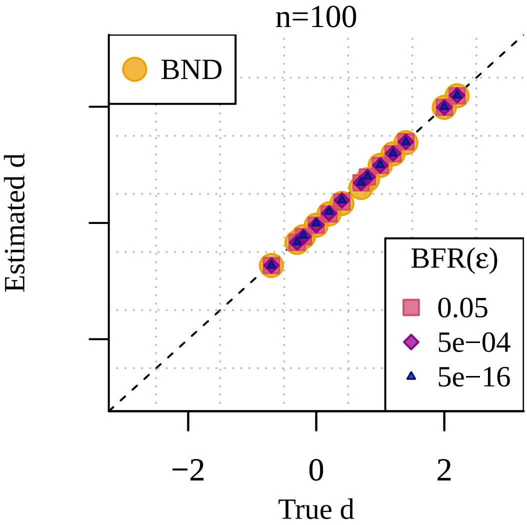
- Likelihood is decreasing in d near boundary $\bar{d}$
- Approximate $(1-\epsilon)\,\%$ quantile of $\hat{d}_{\bar{d}}$ is $< \bar{d}$

$\Pr(\hat{d}_{\bar{d}} > \bar{d}) \approx \epsilon$, small ϵ facilitates:

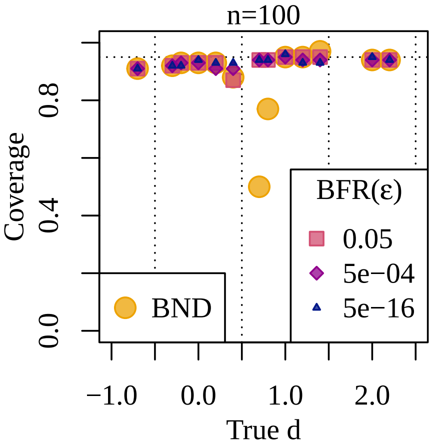
- Construction of CI's for d with nominal coverage
- Implementation of tests of $d = c$ with nominal level

When $\epsilon = 0.5$ we'll call estimator **BND**
When $\epsilon < 0.5$, we'll call estimator $\mathbf{BFR}(\epsilon)$

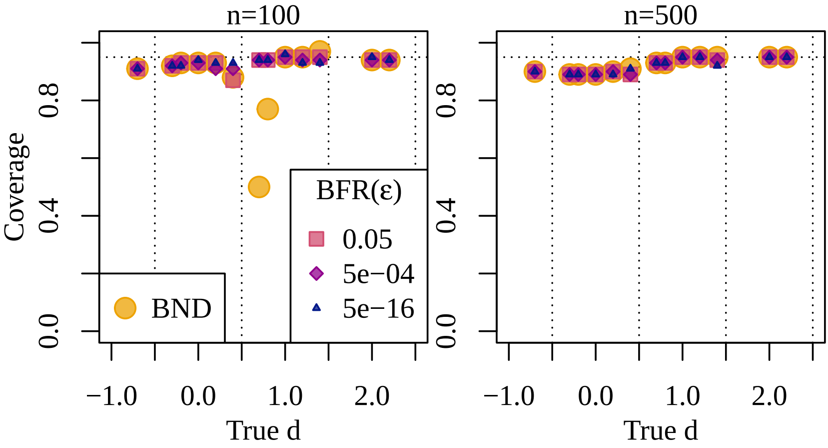
Adaptive Procedure Simulations



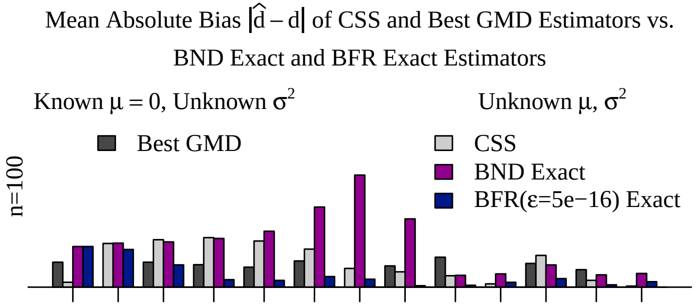
Adaptive Procedure Simulations



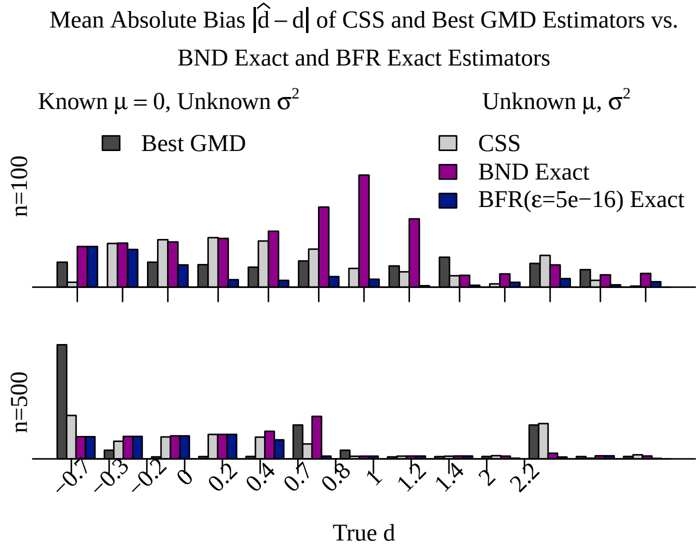
Adaptive Procedure Simulations



Adaptive Procedure vs. Competition

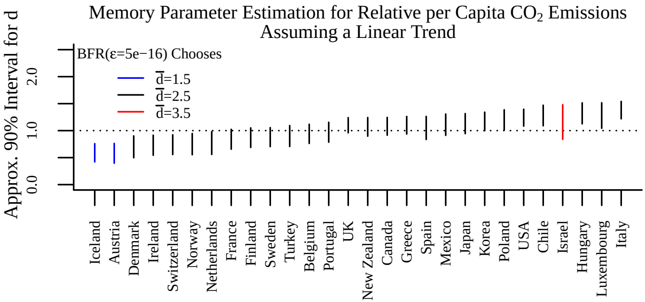


Adaptive Procedure vs. Competition



Application to CO^2 Emission Convergence

$$(1-B)^{d_c}\left(y_{tc}-\mu_c-\beta_c t\right)=z_{tc}, \quad z_{tc} \overset{i.i.d.}{\sim} \text{normal}\left(0,\sigma_c^2\right),$$



Barassi et al. (2018)

Conclusions and Future Work

Conclusions and Future Work

We took a difficult problem and turned it into an easy problem!

Conclusions and Future Work

We took a difficult problem and turned it into an easy problem!

Maximizing over possibly non-stationary FI models
$\approx$ Maximizing a sequence of stationary ARFIMA models

Conclusions and Future Work

We took a difficult problem and turned it into an easy problem!

Maximizing over possibly non-stationary FI models
$\approx$ Maximizing a sequence of stationary ARFIMA models

Everything we have done:

Conclusions and Future Work

We took a difficult problem and turned it into an easy problem!

Maximizing over possibly non-stationary FI models
$\approx$ Maximizing a sequence of stationary ARFIMA models

Everything we have done:

- Extends to ARFIMA(p, d, q)

$$\phi\left(B\right)\left(1-B\right)^{d}\left(y_t-\mu_t\right)=\theta\left(B\right)z_t,\ z_t \overset{i.i.d.}{\sim} \text{normal}\left(0,\sigma^2\right)$$

  - ⋆ Optimization is harder
  - ⋆ Likelihood may be multimodal

Conclusions and Future Work

We took a difficult problem and turned it into an easy problem!

Maximizing over possibly non-stationary FI models
$\approx$ Maximizing a sequence of stationary ARFIMA models

Everything we have done:

- Extends to ARFIMA(p, d, q)

$$\phi(B)(1-B)^d(y_t - \mu_t) = \theta(B) z_t,\ z_t \overset{i.i.d.}{\sim} \text{normal}\left(0, \sigma^2\right)$$

  - ⋆ Optimization is harder
  - ⋆ Likelihood may be multimodal

- Can be made approximate

Baillie, R. T. (1996). Long memory processes and fractional integration in econometrics. Journal of Econometrics 73, 5–59.

Barassi, M. R., N. Spagnolo, and Y. Zhao (2018). Fractional Integration Versus Structural Change: Testing the Convergence of CO_2 Emissions. Environmental and Resource Economics 71(4), 923–968.

Beran, J. (1995). Maximum Likelihood Estimation of the Differencing Parameter for Invertible Short and Long Memory Autoregressive Integrated Moving Average Models. Journal of the Royal Statistical Society. Series B (Methodological) 57(4), 659–672.

Box, G. E. P. and G. M. Jenkins (1970). Time Series Analysis: Forecasting and Control. San Francisco: Holden Day.

Doornik, J. A. and M. Ooms (2003). Computational aspects of maximum likelihood estimation of autoregressive fractionally integrated moving average models. Computational Statistics and Data Analysis 42(3), 333–348.

Granger, C. W. and R. Joyeux (1980). An introduction to long-memory time series models and fractional differencing. Journal of Time Series Analysis 1(1), 15–29.

Hosking, J. R. M. (1981). Fractional Differencing. Biometrika 68(1), 165–176.

Laboissiere, R., J. C. Letievant, E. Ionescu, P. A. Barraud, M. Mazzuca, and C. Cian (2015). Relationship between spectral characteristics of spontaneous postural sway and motion sickness susceptibility. PLoS ONE 10(12), 1–15.

Lo, C.-M., C. R. Keese, and I. Giaever (1993). Monitoring Motion of Confluent Cells in Tissue Culture. Experimental Cell Research 204, 102–109.

Mayoral, L. (2007). Minimum distance estimation of stationary and non-stationary ARFIMA processes. Econometrics Journal 10(1), 124–148.

Odaki, M. (1993). On the Invertibility of Fractionally Differenced ARIMA Processes. Biometrika 80(3), 703–709.

Velasco, C. and P. M. Robinson (2000). Whittle Pseudo-Maximum Likelihood Estimation for Nonstationary Time Series. Journal of the American Statistical Association 95(452), 1229–1243.

Acknowledgements

Thank you to:

- Ruey, for teaching David so he could teach me!

Acknowledgements

Thank you to:

- Ruey, for teaching David so he could teach me!
- My coauthors, David and Gena

Acknowledgements

Thank you to:

- Ruey, for teaching David so he could teach me!
- My coauthors, David and Gena
- Grant support:
    - Xerox PARC Faculty Research Award
    - National Science Foundation Awards 1455172, 1934985, 1940124, and 1940276
    - USAID
    - Cornell University Atkinson Center for a Sustainable Future

Acknowledgements

Thank you to:

- Ruey, for teaching David so he could teach me!
- My coauthors, David and Gena
- Grant support:
    - Xerox PARC Faculty Research Award
    - National Science Foundation Awards 1455172, 1934985, 1940124, and 1940276
    - USAID
    - Cornell University Atkinson Center for a Sustainable Future

Conclusions and Future Work

We took a difficult problem and turned it into an easy problem!

Maximizing over possibly non-stationary FI models
$\approx$ Maximizing a sequence of stationary ARFIMA models

Everything we have done:

- Extends to ARFIMA(p, d, q)

  $$\phi(B)(1-B)^d(y_t - \mu_t) = \theta(B) z_t,\ z_t \overset{i.i.d.}{\sim} \text{normal}\left(0, \sigma^2\right)$$

  - Optimization is harder
  - Likelihood may be multimodal
- Can be made approximate