Time Series Advanced

EC420 MSU

Justin Kirkpatrick

Last updated April 12, 2021

Time series

- Define
- Lag models
- Can we use MLR1-MLR6 and call a time series OLS estimate unbiased?
    - No.

Time series

Stationary vs. Non-Stationary

- Why do we care?

Weak dependence

Some useful time series models:

- Finite Distributed Lag (FDL)
  - Covered last week
- Moving Average (MA) models
- Autoregressive (AR) models

Stationary Time Series

A stationary time series $\{x_t : t = 1, 2, \cdots\}$ is one where the *joint* distribution of (x_t, x_{t+h}) is the same as the joint distribution of $(x_{t+k}, x_{t+h+k}) \forall k, h \geq 1$

Remember the idea of a time series being a realization of a sequence? "Pulling a chain" out of a bag of all possible chains?

Joint Distribution includes

- The mean of x_t
- The mean of x_{t+h}
- The variance of x_t
- The variance of x_{t+h}
- The covariance of (x_t, x_{t+h})

Stationary implies *identically distributed*. Imagine $h = 1$. Then x_t and x_{t+1} have the same mean and variance when stationary.

But **stationary** means even more.

Stationarity says that the relationship over time is stable.

That whatever process drives the relationship between x_t and x_{t+h} also applies, in a random sense, to x_{t+k} and x_{t+h+k}.

Non-stationarity

- Non-stationarity is not uncommon. Think about our time trend regression:

$$y_t = \beta_0 + \beta_1 x_{1,t} + \beta_2 t + u_t$$

  Ignore the $x_{1,t}$ for a moment.

Because of the $\beta_2 t$ term, it is clear that the joint distribution of (y_t, y_{t+1}) is not the same as (y_{t+h}, y_{t+1+h})

- For starters, the mean of y_{t+1} is β_2 higher that y_t, so the joint distribution of (y_t, y_{t+h}) is different from the joint distribution of (y_{t+1}, y_{t+1+h}).
  - But we *can* control for that.
  - When we can't, though, it becomes a problem

Why is non-stationarity a problem?

> (I)f we want to understand the relationship between tow or more variables using regression analysis, we need to asssume some sort of stability over time. If we allow the relationship between two variables to change arbitrarily in each time period, then we cannot hope to learn much about how a change in one affects the other(...)

Much of time series econometrics is about being very specific as to how big of a problem this may be, and when it stops being a problem

- If we assume that everything past one lag is uncorrelated, time series is very easy!
- We already saw that assuming y_t has no effect on x_{t+1} made things very easy!

A weaker form of stationarity is **covariance stationarity**

This holds when:

1. $E[x_t]$ is constant
2. $Var(x_t)$ is constant
3. For any t, h, $Cov(x_t, x_{t+h})$ depends only on h, not on t.

The first two are straightforward. The third simply means that the correlation structure is the same.

- x_1 and x_3 may be correlated...
- ...but x_2 and x_4 have the same correlation
- Correlation is a *population* concept.

This is sometimes called *weak stationarity*

TS5, "no correlation in u_t, u_{t+h}", meets this

Let's pause for a moment

These conditions are all ways of saying that

> "regardless of *where* in the chain our sample is drawn, we can learn about *how* the chain behaves from our observation".

Any questions?

Stationarity is about how stable the relationship is between (x_t, x_{t+h})

For a variety of t and h

We need a concept that tells us how large h has to be to say that x_t and x_{t+h} are essentially unrelated

Weak dependence

A **weakly dependent** time series $\{x_t : 1, 2, \cdots\}$ is one where, as h gets larger, x_t and x_{t+h} become "almost independent".

- If it is *covariance stationary* and
- $Cov(x_t, x_{t+h}) \to 0$ as $h \to \infty$
- "Asymptotically uncorrelated"

MICHIGAN STATE UNIVERSITY

If a time series is weakly dependent, then we have a Central Limit Theorem (CLT) and Law of Large Numbers (LLN) that can apply

The CLT is what let us say that we could ignore non-normal errors The LLN is what let us say that averages, with large enough N, are unbiased estimates of the population average.

- This let us say that $E[\hat{\beta}] = \beta$

One useful weakly dependent model: **Moving Average**

$$x_t = e_t + \alpha_1 e_{t-1}$$

- e_t is a random i.i.d. sequence with zero mean and constant variance
- The process $\{x_t\}$ is a **Moving Average of order one (MA(1))**

Stationary TS

$$x_t = e_t + \alpha_1 e_{t-1}$$

t	e	x
1	-0.2	NA
2	-1.4	-1.50
3	0.4	-0.30
4	1.5	1.70
5	0.0	0.75

x has constant mean, constant variance

x_t and x_{t+h} have the same covariance $\forall\, h$

We can go beyond the first lag of e and have a MA(2) or more...

$$x_t = e_t + \alpha_1 e_{t-1} + \alpha_2 e_{t-2}$$

- This is still stationary
- This is still weakly dependent

Another useful model: **Autoregressive**

$$y_t = \rho_1 y_{t-1} + e_t$$

- $\{e_t : t = 1, 2, \cdots\}$ is i.i.d. with zero mean and constant variance σ_e^2
- So e_t is independent of y
- Each y_t is equal to some fraction of the previous y_t **and** that new error term, e_t .
    - Sometimes called an "innovation"
- There is some y_0 that started it all

Imagine for a moment that $\rho_1 >> 1$

- How does y_t behave over time?

The **AR(1)** process:

$$y_t = \rho_1 y_{t-1} + e_t$$

We can show that, if $\rho_1 < 1$, then y_t, our **AR(1)** process, is weakly dependent.

- $\rho_1 < 1$ is called a "stable AR(1)" This is largely because $e_t \perp y_{t-1}$ (independent)

$Var(y_t) = Var(\rho_1 y_{t-1}) + Var(e_t) = \rho_1^2 Var(y_{t-1}) + Var(e_t)$

$\Rightarrow \sigma_y^2 = \rho^2 \sigma_y^2 + \sigma_e^2$

As long as $\rho < 1 \rightarrow \rho^2 < 1$, then we can get:

$$\sigma_y^2 = \frac{\sigma_e^2}{1 - \rho_1^2}$$

Wooldridge 11-1 shows that

$Corr(y_t, y_{t+10}) > Corr(y_t, y_{t+20})$ when $\rho < 1$

Thus, AR(1) is weakly dependent.

But how does weakly dependent help?

Asymptotic Properties of TS OLS

New Assumptions

TS.1' - Linearity and Weak Dependence

$$\{(\mathbf{x}_t, y_t) : t = 1, 2, \cdots\} \text{ is stationary and weakly dependent}$$

in a model such as:

$$y_t = \beta_0 + \beta_1 x_{1,t} + \beta_2 x_{1,t-1} + \cdots + \beta_j y_{t-1} + u_t$$

Note that

- $\mathbf{x}$ has x and lags of x
- $\mathbf{x}$ also has lags of y

Really, we just need weak dependence, but Wooldridge includes stationarity

TS.2' - No Perfect Multicolinearity (same as before)

New Assumptions

TS.3' - $\mathbf{x}_t$ is contemporaneously exogenous

$$E[u_t|\mathbf{x}_t] = 0$$

We've relaxed that pernicious **strict exogeneity** assumption from before by adding TS.1'

- Remember, **strict exogeneity** was $E[u_t|\mathbf{X}] = 0$, all future values of $\mathbf{x}_t$.
- So as long as TS.1' and TS.3' hold, we don't have to worry about correlation between u_{t-1} and x_t, even when it's because x_t is related to past y_{t-1}!

New Assumptions

With TS.1', TS.2', TS.3', $\hat{\beta}$ is consistent

Not necessarily unbiased, but with larger and larger N, the estimate gets better!

So our **AR(q)** model estimators, if it meets TS.1'-TS.3', are consistent.

- To meet TS.1', it has to be stable

So are our **MA(q)** model estimators.

So are our FDL estimators, even when future values of x_t are affected by y_{t-1}

TS.4' - Contemporaneous homoskedasticity

$$Var(u_t|\mathbf{x}_t) = \sigma^2$$

As opposed to TS.4 - $Var(u_t|\mathbf{X}) = \sigma^2$

TS.5' - No Serial Correlation

$$Corr(u_t, u_s|\mathbf{x}_t, \mathbf{x}_s) = 0 \quad \forall t, s$$

As Wooldridge says, ignore the conditioning and just think of whether or not u_t and u_s are correlated.

Note that specifying an **AR(1)** when the real process is **AR(2)** results in serial correlation (because y_{t-2} is in the error and y_t is serially correlated)

Under TS.1'-TS.5', the errors are asymptotically normally distributed

Which lets us use t-statistics, F-tests, confidence intervals, p-value, etc.

We've skipped assuming normal errors and gone straight to using the **asymptotic properties**.

So we need to be able to say we have a large N

- Which means a large T

Serial correlation in u is not the end of the world. We won't get to the solution, but briefly, here's how it works:

Heteroskedasticity and Autorcorrelation-Consistent Errors

or **HAC**, are calculated for each β_j by multiplying each "naive" OLS std. error by a correction factor, $\hat{v}_j$

- $\hat{v}_j$ is a function of two things:
    - u_t, as one would expect
    - r_t, where r_t is the error in $x_{jt} = \alpha_0 + \alpha_1 x_{kt} + \cdots + r_t$
    - When $\hat{u}_t$ covaries with $\hat{r}_t$ (the part of x_{jt} not explained by $\mathbf{x}_t$), the correction factor $\hat{v}_j$ gets bigger.
    - When lags of $\hat{u}_t \times \hat{r}_t$ covary, $\hat{v}_j$ gets bigger.
    - How many lags to include is a question for another day (1? 2? 10?)

Random Walk

MICHIGAN STATE UNIVERSITY

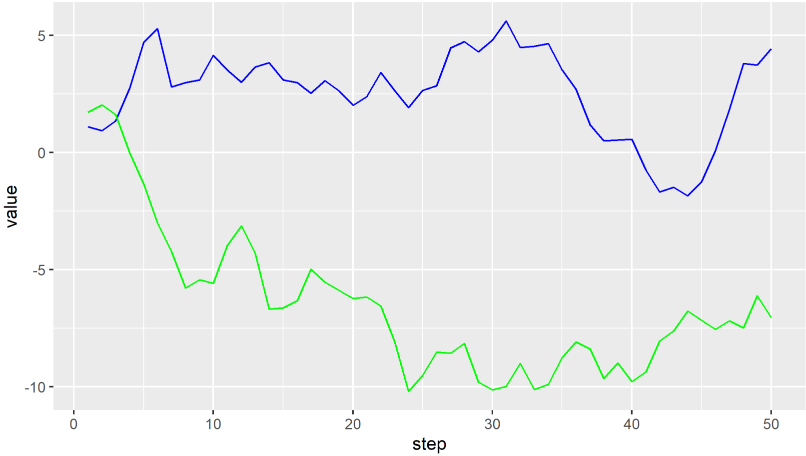
A **random walk** is a highly persistent time series

It is not *weakly dependent*, so it poses a problem to TS.1'-TS.5'

A random walk is a process that follows:

$$y_t = \rho y_{t-1} + e_t \text{ where } \rho = 1$$

and e is iid, $E[e] = 0$ (mean zero errors) and $Var(e) = \sigma_e^2$



We can write

$$y_t = y_{t-1} + e_t$$

which is the same as

$$y_t = y_{t-2} + e_{t-1} + e_t$$

which generalizes to:

$$y_t = y_0 + e_1 + e_2 + \cdots + e_t$$

A Random Walk has a constant $E[y_t]$

$$y_t = y_0 + e_1 + e_2 + \cdots + e_t$$

means

$$E[y_t] = E[y_0] + 0 + 0 + \cdots + 0$$

But a Random Walk does not have constant variance:

$$Var(y_t) = Var(y_0) + Var(e_1) + Var(e_2) + \cdots + Var(e_t)$$

Which is $t\sigma_e^2$, so it changes over time!

A Random Walk is not stationary and TSR.1'-TSR.5' do not hold!

Nor is a Random Walk covariance-stationary (depends on t), and is not weakly dependent.