EC402 - Problem Set 7

Konrad Burchardi

18th of March 2009

Introduction

Today we will

- talk about what ‘weak stationarity’ is,
- see how to estimate the parameter of an MA(1) process,

And we will

- simulate the distribution of the t-statistic when the data series has a unit-root and
- see how it is useless.

Question 1.A

Weak Stationarity

Question: Is

$$y_t = \Sigma_{j=0}^{\infty} \theta_j \epsilon_{t-j} = \theta_0 \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + ... \qquad (1)$$

with $\Sigma_{j=0}^{\infty} |\theta_j| < \infty$ and $\{\epsilon_t\}$ a m.d.s., $E[\epsilon_t] = 0$ and $E[\epsilon^2] = \sigma^2$, **weakly stationary**?

Definition (Weak Stationarity)

A time-series process, $\{z_t\}_{t=-\infty}^{t=\infty}$, is weakly stationary if

> $E[z_t]$ *is finite and independent of* t *and*
>
> $Cov(z_t, z_{t-k})$ *is finite and independent of* t.

Definition (Martingale Difference Sequence)

A sequence $\{z_t\}$ *is a martingale difference sequence if* $E[z_t|z_{t-1}, z_{t-2}, ...] = 0$.

Question 1.A
Weak Stationarity

Check whether conditions for weak stationarity is satisfied:

1.) $E[y_t] = \theta_0 E[\epsilon_t] + \theta_1 E[\epsilon_{t-1}] + \theta_2 E[\epsilon_{t-2}] + ... = 0$

2.) $Cov(y_t, y_{t-k}) = E[y_t y_{t-k}] = E[(\theta_0 \epsilon_t + \theta_1 \epsilon_{t-1} + ...)(\theta_0 \epsilon_{t-k} + \theta_1 \epsilon_{t-k-1} + ...)]$

Note: Because ϵ_t is a m.d.s. by definition $E[\epsilon_t | \epsilon_{t-1}, \epsilon_{t-2}, ...] = 0$. Since mean-independence implies zero-covariance and since $E[\epsilon_t] = 0$ we know $E[\epsilon_i \epsilon_j] = Cov(\epsilon_i \epsilon_j) = 0$ for all $i \neq j$. Therefore if we open up the brackets in the above formula and take expectations all terms where the ϵ's have different subscripts drop out. Hence, using $E[\epsilon_t^2] = \sigma^2$

$$Cov(y_t, y_{t-k}) = \theta_k \theta_0 E[\epsilon_{t-k}^2] + \theta_{k+1} \theta_1 E[\epsilon_{t-k-1}^2] + ... = \sigma^2 \Sigma_{p=0}^{\infty} \theta_{k+p} \theta_p$$

Are both **finite** and **independent of** t?

Question 1.A

Weak Stationarity

Only thing we need to show is that $\gamma_k \equiv Cov(y_t, y_{t-k})$ is finite for all k or

$$|\gamma_k| < \infty$$

Proof: Note that $|a + b| \leq |a| + |b|$ and hence

$$|\gamma_k| \leq \sigma^2 \Sigma_{p=0}^{\infty} |\theta_{k+p}\theta_p|$$

Now, if we can show that $\Sigma_{k=0}^{\infty}|\gamma_k| < \infty$ this implies that for every k we have $|\gamma_k| < \infty$. Summing the above and noting that $|\theta_{k+p}\theta_p| = |\theta_{k+p}| \cdot |\theta_p|$ gives

$$\Sigma_{k=0}^{\infty}|\gamma_k| \leq \sigma^2 \Sigma_{k=0}^{\infty}\Sigma_{p=0}^{\infty}|\theta_{k+p}| \cdot |\theta_p| = \sigma^2 \Sigma_{p=0}^{\infty}|\theta_p|\Sigma_{k=0}^{\infty}|\theta_{k+p}|$$

But since by assumption there is a M such that $\Sigma_{p=0}^{\infty}|\theta_p| = M < \infty$ surely for every k we have $\Sigma_{k=0}^{\infty}|\theta_{k+p}| \leq M$ and hence

$$\Sigma_{k=0}^{\infty}|\gamma_k| \leq \sigma^2 \Sigma_{p=0}^{\infty}|\theta_p|\Sigma_{k=0}^{\infty}|\theta_{k+p}| \leq \sigma^2 M \Sigma_{p=0}^{\infty}|\theta_p| = \sigma^2 M^2 < \infty$$

Question 1.B/C

What are we up to?

Question: How to estimate θ in

$$y_t = \theta\epsilon_{t-1} + \epsilon_t \tag{2}$$

given that we do not know ϵ_{t-1}?

Idea: We can find a consistent estimate, $\hat{u}_t$, of the ϵ_t's (step 1; Q2.B) and then run the above regression using this $\hat{u}_{t-1}$ (step 2; Q2.C).

Question 1.B

How to find a consistent estimate of ϵ_t?

We know that the true DGP is $y_t = \theta\epsilon_{t-1} + \epsilon_t = (1 + L\theta)\epsilon_t$. We know that if $|\theta| < 1$ this is 'invertible', so we can write it as the $AR(\infty)$

$$\begin{aligned}\epsilon_t &= \frac{y_t}{1 + L\theta} \\ &= y_t + (-\theta)y_{t-1} + ... + (-\theta)^p y_{t-p} + \Sigma_{j=p+1}^{\infty}(-\theta)^j y_{t-j}\end{aligned}$$

or, as well it will be true,

$$y_t = -(-\theta)y_{t-1} - ... - (-\theta)^p y_{t-p} - \Sigma_{j=p+1}^{\infty}(-\theta)^j y_{t-j} + \epsilon_t$$

The y_t we know, so this we can estimate! And since this is a stationary, ergodic sequence, we know that when we run the regression of y_t on past values of y up to $t - p$ the coefficients ρ are consistent[1], so

$$plim(\hat{\rho}_i) = (-1)^{i+1}\theta^i$$

Remember, we wanted to find a consistent estimate of ϵ_t. So what seems like a promising way to get this?

---

[1] You might ask: Why don't we stop here, we already have a consistent estimate of θ? The answer seems to be that either we could not get correct standard errors on this estimate or the proposed procedure is more efficient.

Question 1.B

How to find a consistent estimate of ϵ_t?

We just calculate

$$\hat{u}_t = y_t - \hat{\rho}_1 y_{t-1} - ... - \hat{\rho}_p y_{t-p} \tag{3}$$

But is this consistent? Plug in for y_t and check:

$$\begin{aligned} plim(\hat{u}_t) &= plim(y_t - \hat{\rho}_1 y_{t-1} - ... - \hat{\rho}_p y_{t-p}) \\ &= plim([\theta - \hat{\rho}_1]y_{t-1} + ... + [-(-\theta)^p - \hat{\rho}_p]y_{t-p} - \Sigma_{j=p+1}^{\infty}(-\theta)^j y_{t-j} + \epsilon_t) \end{aligned}$$

Since we saw before that $plim[\theta - \hat{\rho}_1] = 0$ and similarly for the first expressions we find

$$plim(\hat{u}_t) = plim(\underbrace{-\Sigma_{j=p+1}^{\infty}(-\theta)^j y_{t-j}}_{\omega_t} + \epsilon_t)$$

Question 1.B

How to find a consistent estimate of ϵ_t?

We can rewrite

$$\begin{aligned} \omega_t &= -\Sigma_{j=p+1}^{\infty}(-\theta)^j y_{t-j} \\ &= -(-\theta)^{p+1}\Sigma_{j=0}^{\infty}(-\theta)^j y_{t-j-p-1} \\ &= -(-\theta)^{p+1}(\epsilon_{t-p} + \theta\epsilon_{t-p+1}) \end{aligned}$$

So, as $p \to \infty$, $\omega_t \to 0$. Hence, since ϵ_t is the true value

$$plim(\hat{u}_t) = plim(\epsilon_t) = \epsilon_t$$

Question 1.C

How to find a consistent estimate of θ?

In the second step of the procedure we estimate θ in the regression

$$y_t = \theta \hat{u}_{t-1} + \nu_t \tag{4}$$

Question: Will the $\hat{\theta}_{OLS}$ from this be consistent?

Question 1.C

How to find a consistent estimate of θ?

The OLS estimate will be

$$\hat{\theta}_{OLS} = \frac{\Sigma \hat{u}_{t-1} y_t}{\Sigma \hat{u}_{t-1}^2} = \frac{\frac{1}{T}\Sigma \hat{u}_{t-1}(\epsilon_t + \theta\epsilon_{t-1})}{\frac{1}{T}\Sigma \hat{u}_{t-1}^2} = \frac{\frac{1}{T}\Sigma \hat{u}_{t-1}\epsilon_t}{\frac{1}{T}\Sigma \hat{u}_{t-1}^2} + \theta \frac{\frac{1}{T}\Sigma \hat{u}_{t-1}\epsilon_{t-1}}{\frac{1}{T}\Sigma \hat{u}_{t-1}^2} \tag{5}$$

We can show that

$plim(\frac{1}{T}\Sigma \hat{u}_{t-1}^2) = \sigma^2$

$plim(\frac{1}{T}\Sigma \hat{u}_{t-1}\epsilon_{t-1}) = n$, with $n \to 0$ as $p \to \infty$

$plim(\frac{1}{T}\Sigma \hat{u}_{t-1}\epsilon_t) = \theta\sigma^2 + m$, with $m \to 0$ as $p \to \infty$

Hence: $plim(\hat{\theta}_{OLS}) = \theta$ as $p \to \infty$

Question 1.D

How does the precision of this depend on p?

Question: How does the precision of this depend on p?

- In large samples, the quality of the first step approximation improves as p goes to infinity. This is because, the larger is p, the smaller is our residual term ω_t. In addition, the variance of the $\hat{u}_t$ is declining in p.
- However, in small samples, as p goes to infinity, the standard errors for our ρ's grow. Intuitively, this is because we are using smaller and smaller samples as we allow the specified lag lengths to grow.[2]

[2] It turns out (you did not need to know this for this question but it may be a useful thing to know) the optimal rate for $\rho \to \infty$ is such that $p^2/T \to 0$.

Question 2

Asymptotic distribution of our test statistic

Remember from Vassilis that we derived the exact distributions of the **F**- and **t**-statistic under **A1**-**A5N**.

If we do not have **A5N** then we could not derive the finite sample distribution, but we could show that asymptotically the finite sample distribution is still correct. However, for this proof (which we did not do in detail) we **required that the data involved is weakly stationary**!

Question 2

Asymptotic distribution of our test statistic

Consider the simple case

$$y_t = \beta y_{t-1} + \epsilon_t \tag{6}$$

with $\epsilon_t \sim$ i.i.d. $N(0,1)$ and $y_0 = 0$.

What will the asymptotic distribution look like if

(i) $\beta_1 = 1.0$

(ii) $\beta_2 = 0.9$

(iii) $\beta_3 = 0.5$

Question 2

Monte Carlo experiment

Answer: We do not know exactly, but we can simulate it!

Idea: Given a true parameter, we just create a random data series with this parameter and calculate the **t**-statistic. Then we do this 1000 times and see how the **t**-statistic is distributed.

Question 2

Monte Carlo experiment - Stata code

In each of the 1000 repetition we start with generating random data

```
clear

set seed 23`i'456
set obs 500
gen e = invnorm(uniform())  <----- generate random epsilon's

gen t = 1
replace t = 1+ [_n-1]       <----- give time subscripts
tsset t

gen b1 = 1.0                <----- let's do it first for this beta
```

Question 2

Monte Carlo experiment - Stata code

Using a $y_0 = 0$, the ϵ_t's generated and β_1 we can calculate the random series y_t:

```
for num 1/3: gen yX = 0 if t ==0        <----- Set's y_0=0
for num 1/3: replace yX = e if t == 1   <----- Set's y_1=\epsilon_t

replace y1 = b1*y1[_n-1]+ e if t>1      <----- Calculates rest of series

for num 1/3: gen yX_lag = yX[_n-1]      <----- Creates for each series
                                               one of lagged values
```

Question 2

Monte Carlo experiment - Stata code

Then we run a regression and see what the **t**-statistic would be:

```
reg y1 y1_lag
gen b1_ols = _b[y1_lag]
gen mistake1 = b1_ols - b1
gen seb1 = _se[y1_lag]
gen tstat1 = mistake1/seb1
```

And save it:

```
gen ones = 1
collapse b1_ols mistake1 tstat1, by(ones)
save model`i', replace
```

We just repeat this 1000 times with

```
local i = 1
while `i' <= 1000 {

...

local i = `i' + 1
}
```

and then put all the saved coefficient estimates, mistakes and t-statistics together

```
use model1, clear
for num 2/1000: qui append using modelX
```

Now we can see

- how far we were off on average (mean of mistake),
  `sum mistake1`
- plot the simulated distribution of the OLS estimator and
  ```
  kdensity b1_ols
  graph export b1.png
  ```
- plot the simulated distribution of the **t**-statistic
  ```
  kdensity tstat1
  graph export tstat1.png.
  ```

We would then repeat this for $\beta = 0.9$ and $\beta = 1.0$.[3]

[3] If you look at Radha's code, do not get confused. The only difference is that every time she creates a series of ϵ's she uses this to calculate the y_t for all 3 values of β not to repeat the creation of random data.

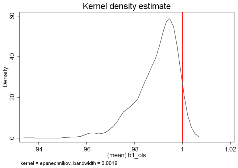
Simulated **bias**:

Variable	Obs	Mean	Std. Dev.	Min	Max
bias1	1000	-.0107294	.0088402	-.0644693	.0048755
bias2	1000	-.0073681	.0202228	-.0947172	.0462514
bias3	1000	-.003086	.0389725	-.1618057	.1083517

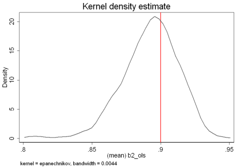
Question 2

Monte Carlo experiment - Results

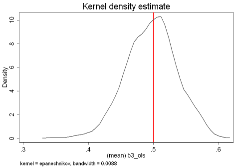
Simulated **distribution of OLS estimator**:



(a) $\beta = 1.0$



(b) $\beta = 0.9$

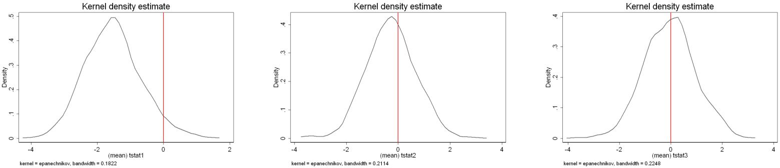


(c) $\beta = 0.5$

Question 2

Monte Carlo experiment - Results

Simulated **distribution of the t-statistic**:



(a) $\beta = 1.0$ (b) $\beta = 0.9$ (c) $\beta = 0.5$