AR(1) and unit-root processes

William Mann

EMORY
GOIZUETA BUSINESS SCHOOL

AR(1) process and random-walk process

AR(1) means "autoregressive process of order 1". It looks like this:

$$y_t \quad = \quad \mu \quad + \quad \phi y_{t-1} \quad + \quad u_t$$

where each u_t is an independent draw from $\mathrm{N}(0, \sigma^2)$.

An AR(1) process eventually stabilizes around an average, $\frac{\mu}{1-\phi}$.

If $\phi = 1$, we get a **random walk**, an example of a **unit-root** process:

$$y_t \quad = \quad \mu \quad + \quad y_{t-1} \quad + \quad u_t$$

This process does *not* have any long-run "average."

- This is a very common case with financial data.
- It also creates some challenges for statistical modeling!
  To understand why, we need the idea of a "stationary" process...

Weakly stationary process (or covariance stationary)

A process y is **weakly stationary** (or **covariance stationary**) if

- $\mathbb{E}[y_t]$ does not depend on t,
- $Var(y_t)$ does not depend on t,
- $\text{Cov}(y_t, y_s)$ depends only on the distance $t - s$.

In words, this means that y follows a stable distribution at all times.

From now on, I will use "stationary" as a shorthand for "weakly stationary," and will use "nonstationary" to mean "not weakly stationary,"

Other fields outside economics sometimes use these terms differently.

Stationary and nonstationary examples

An AR(1) with $\phi < 1$ is stationary:

$$\begin{aligned} \mathbb{E}[y_t] &= \frac{\mu}{1-\phi} \\ Var(y_t) &= \frac{\sigma^2}{1-\phi^2} \\ Cov(y_t, y_s) &= \phi^{t-s} \times \frac{\sigma^2}{1-\phi^2} \end{aligned}$$

A random walk is *not* stationary:

$$\begin{aligned} \mathbb{E}[y_t] &= t \times \mu \quad + \quad y_0 \\ Var(y_t) &= t \times \sigma^2 \\ \mathrm{Cov}(y_t, y_s) &= \min(t, s) \times \sigma^2 \end{aligned}$$

The difference of a random walk is stationary

The *difference* of a random walk *is* stationary.

To see this, use Δy_t to represent $y_t - y_{t-1}$. Then we can write

$$\Delta y_t \quad = \quad \mu + u_t$$

and from this we can see that

$$\begin{aligned} \mathbb{E}[\Delta y_t] &= \mu \\ Var(\Delta y_t) &= \sigma^2 \\ \text{Cov}(\Delta y_t, \Delta y_s) &= 0 \end{aligned}$$

This is an important fact that we will use a lot.

Why is stationary data important?

- Statistics is all about *estimating* averages, variances, etc.
  This implicitly assumes that such numbers actually exist!
- For a nonstationary process, they do not.
  So the results of such analysis would not be meaningful.
  We have to find some other approach instead.
- What happens if we (mistakenly) ignore this issue,
  and perform our usual analysis using nonstationary data?
- The computer will not stop us.
  It will still run the analysis and give us (meaningless) results.
- Instead, we have to check for nonstationary data ourselves, and figure out what to do about it. This is our next topic.

Testing for unit roots

Types of nonstationary process

- Statistical tests require a precise null and alternative hypothesis.
- In theory, there are many ways a process can be nonstationary. Each type of nonstationary behavior requires its own test.
- In practice, the most plausible and common category of nonstationary processes is called *unit root* processes.
  - A unit-root process is one that is not stationary on its own, but can be turned into a stationary process through differencing.
  - As we saw earlier, a random walk process is one example.
- So we will focus our attention on tests for unit-root behavior.

Estimating AR(1) and AR(p) processes

Unit root tests build on the following information:

Recall the definition of an AR(1) process,

$$y_t \quad = \quad \mu \quad + \quad \phi y_{t-1} \quad + \quad u_t$$

Suppose you believe a data series follows an AR(1) process.
How would you estimate ϕ? Just regress y_t on y_{t-1}.

An AR(p) process is like an AR(1), with p lags of y:

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + ... + \phi_p y_{t-p} + u_t$$

You can estimate all the ϕ at once, by running a multiple regression of y_t on $y_{t-1}, \ldots, y_{t-p}$. See notebook on Canvas for examples.

Unit root tests: Dickey-Fuller

The Dickey-Fuller unit root test assumes the data are AR(1), and tests the null $\phi = 1$ against the alternative $\phi < 1$.

- Write out Δy_t, which is stationary under either hypothesis:

$$\Delta y_t \quad = \quad \mu \quad + \quad (\phi - 1) \times y_{t-1} \quad + \quad u_t$$

- Estimate $\phi - 1$ by regressing Δy_t on y_{t-1}.
  Then we want to test whether this coefficient is negative.
  If the p-value is low enough, we can reject that $\phi = 1$.
- The test statistic is the t-stat from the regression.
  But, the typical regression p-values will not be correct:
  Sampling distribution is nonstandard under the null $\phi = 1$.
- Refer to special tables of critical values to check significance.

Unit root tests: Augmented Dickey-Fuller (ADF)

Same approach, but control for lags of Δy_t in the regression, allowing the process to be any AR(p) instead of just AR(1).
For example, we can allow y to be AR(2) by running

$$\Delta y_t \quad = \quad \mu \quad + \quad (\phi - 1) \times y_{t-1} \quad + \quad \delta \times \Delta y_{t-1} \quad + \quad u_t$$

Each extra lag allows more general dynamics in y.

- How far to go with this? There is no simple rule.
- One common procedure is to select based on information criteria such as AIC or BIC (see definitions next week).
- Most software allows you to automate this process.

In practice, ADF is the most common unit-root test.

Deterministic trend as alternative hypothesis

In the tests that we described just now:

- the null hypothesis was that the process has a unit root,
- the alternative was that the process is stationary.

Another potential behavior is a *deterministic* trend.

- This is a trend that is a simple function of time.
- Also called *trend-stationary* behavior.

This can be specified as the alternative in an ADF test by

- adding a time trend in the regression,
- or, running the test on detrended data.

(In practice, just specify the appropriate option in software.)

However, deterministic trends are considered to be rare.

Unit root as the alternative: KPSS

The KPSS test reverses the ADF procedure:
Stationary data is the null, unit root is the alternative.
Hence we hope to *not* reject the null under this test.

The derivation is a bit too involved to describe here,
but it's implemented in any statistical software.

KPSS and ADF are often used together.
We hope they give consistent answers (one reject, one fail to reject).
But there is no guarantee of this. Both can reject, or fail to reject.

(KPSS = Kwiatkowski, Phillips, Schmidt, and Shin.)

General comments about unit root tests

- Statistical tests might give you the impression that they are definitive and conclusive, but they are not. No test is perfect.
- Unit root tests in particular have low power.
  This means they often struggle to give us any clear conclusion.
- By the same token, these tests do not give us a cookbook that we can follow to tell us exactly how to handle a time series.
- Instead, view them as a way to check intuition that you already have from background knowledge or economic theory.

Deal with unit-root data by differencing

The standard thing to do with unit-root data is to difference it.

- After one or more differencing steps, the data will be stationary. (In all our examples, one step will be enough.)
- Statistical analysis on this differenced data will be valid.

For example, instead of modeling the price of a stock or an index, we model its *returns*, which are the difference of log price.

The major sacrifice is that you give up on modeling y_t directly, and only attempt to explain the change Δy_t.

So, you don't want to difference data unless absolutely necessary.

But if the series has a unit root, you typically have no choice.