# Non-Stationarity

March 16, 2019

The material in this set of notes is based on S&S Chapter 3, Sections 3.7-3.9, and S&S Chapter 5, with the exception of the material on non-stationarity tests based on **ARIMA**$(p, d, q)$ models. This material is based on Section 2.7.5 of Tsay (2010), as well as several published journal articles introducing and reviewing these tests:

- Dickey and Fuller (1981), "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root";
- Said and Dickey (1984), "Testing for unit roots in autoregressive-moving average models of unknown order";
- Said and Dickey (1985), "Hypothesis Testing in **ARIMA**$(p, 1, q)$ Models";
- Phillips and Perron (1988), "Testing for a unit root in time series regression";
- Schwert (1989), "Tests for Unit Roots: A Monte Carlo Investigation".

I don't expect you to all to read them, but I thought they actually offered clearer explanations of what these tests are doing than the textbooks I have looked at so I wanted to share them as useful resources for anyone who is curious.

# Review

Let's think back to when we first introduced the concept of stationarity. We defined a **stationary** time series $x_t$ as having finite second moments, i.e. $\mathbb{E}\left[x_t^2\right] < \infty$ for all $t$, a constant mean function, $\mu_{x,t} = \mu_x$, and an autocovariance function $\kappa_x(s,t)$ that depends on $s$ and $t$ only through their absolute difference $h = |s-t|$. So far, we've just been using stationary models and crossing our fingers and hoping that our data is stationary. Now we'll start thinking about:

- How to assess whether or not a specific observed time series $\boldsymbol{x}$ is stationary;
- What kind of models to use if we conclude that an observed time series $\boldsymbol{x}$ is not stationary.

We'll see that these ideas are related.

A basic first step to assessing stationarity is always to just plot the time series and examine it carefully, assessing whether or not it looks like the mean and variance are constant over the entire time interval. Some further exploratory analysis can be performed by binning the data and examining how the bin means and variances change - if the time series corresponds to a stationary process and each bin contains enough observations, the bin means and variances should all be very similar and should not display any systematic trends.

If we want to take a more sophisticated approach to assessing stationarity, we might want to consider a parametric model for the data, where the parameter values determine whether or not the model is stationary. This will allow us to develop a hypothesis test for non-stationarity. When we do this, we'll tend to consider two different types of non-stationarity (1) non-stationarity of the mean and (2) non-stationarity of the variance.

# Mean Non-Stationarity & The ARIMA$(p, d, q)$ Model

The workhorse (most commonly used) model for data that displays non-stationarity of the mean is the **ARIMA**$(p, d, q)$ model, which generalizes the **ARMA**$(p, q)$ model that we are already familiar with. Let's introduce some new notation first, because we're going use the idea of differencing a time series to define the **ARIMA**$(p, d, q)$ model. We introduce a differencing operator $\nabla^d$ defined according to:

$$\begin{aligned}
\nabla x_t &= x_t - x_{t-1} \\
\nabla^2 x_t &= \nabla x_t - \nabla x_{t-1} = x_t - 2x_{t-1} + x_{t-2} \\
\nabla^3 x_t &= \nabla^2 x_t - \nabla^2 x_{t-1} = x_t - 3x_{t-1} + 3x_{t-2} - x_{t-3} \\
&\vdots \\
\nabla^k x_t &= \nabla^{k-1} x_t - \nabla^{k-1} x_{t-1}.
\end{aligned}$$

Differencing is a very useful concept because it can be used to address certain kinds of mean non-stationarity. We can show this via a few simple examples. Suppose that we observe a nonstationary time series with a linear trend in time

$$x_t = a + bt + w_t,$$

where $w_t \overset{i.i.d.}{\sim} \mathcal{N}\left(0, \sigma_w^2\right)$. What happens if we difference the observed time series? We obtain

$$\begin{aligned}
\nabla x_t =& a + bt + w_t - \left(a + b\left(t - 1\right) + w_{t-1}\right) \\
=& b + w_t - w_{t-1}.
\end{aligned}$$

This is a **MA** $(1)$ process with mean $b$, so we get a stationary process!

What if we'd had an even more complicated trend over time, e.g. a quadratic trend,

$$x_t = a + bt + ct^2 + w_t.$$

What happens if we difference the observed time series in this case? We obtain

$$\begin{aligned}\nabla x_t =& a + bt + ct^2 + w_t - \left(a + b\left(t-1\right) + c\left(t-1\right)^2 + w_{t-1}\right) \\ =& b + 2ct + c + w_t - w_{t-1}.\end{aligned}$$

This isn't stationary yet - we still have a linear trend in time. What if we difference again?

$$\begin{aligned}\nabla^2 x_t =& \nabla x_t - \nabla x_{-1} \\ =& \left(b + 2ct + c + w_t - w_{t-1}\right) - \left(b + 2c\left(t-1\right) + c + w_{t-1} - w_{t-2}\right) \\ =& 2c + w_t - 2w_{t-1} + w_{t-2}.\end{aligned}$$

This is a **MA** $(2)$ process with mean $2c$, so again we get a stationary process!

This leads us to the definition of the **ARIMA**$(p, d, q)$ model. A process $x_t$ is said to be **ARIMA**$(p, d, q)$ if

$$\phi\left(B\right)\left(\nabla^d x_t - \mu_x\right) = \theta\left(B\right) w_t, \tag{1}$$

where $w_t \overset{i.i.d.}{\sim} \mathcal{N}\left(0, \sigma_w^2\right)$ and $\mu_x = \mathbb{E}\left[\nabla^d x_t\right]$. This means that the differenced time series $\nabla^d x_t$ is an **ARMA**$(p, q)$ process. The **ARIMA**$(p, d, q)$ model generalizes the **ARMA**$(p, q)$ model, insofar as setting $d = 0$ yields an **ARMA**$(p, q)$ model. The **ARIMA**$(p, d, q)$ gives us a parametric framework for assessing stationarity - if $d = 0$, $x_t$ is a stationary process, whereas if $d > 0$ $x_t$ is non-stationary. This leads us to the three tests of non-stationarity.

- The **Dickey-Fuller** test tests the null hypothesis $H$ that $x_t$ is an **ARIMA** $(0, 1, 1)$ process against the alternative hypothesis that $x_t$ is a stationary **ARMA** $(1, 0, 0)$ process. It assumes that

  $$\nabla x_t = a + \kappa x_{t-1} + w_t, \tag{2}$$

  where $w_t \sim \overset{i.i.d.}{\sim} \mathcal{N}\left(0, \sigma_w^2\right)$. The null and alternative hypotheses can be equivalently expressed as $\kappa = 0$ and $\kappa \neq 0$, respectively. The test statistic is the $t$-statistic $\hat{\kappa}/\mathrm{se}\left(\hat{\kappa}\right)$ for the the least-squares estimate $\hat{\kappa}$ based on (2), and the approximate distribution of

$\hat{\kappa}$ under the null as $n \to \infty$ is the Dickey-Fuller distribution. This is a non-standard distribution, but the relevant quantiles have been derived. Although useful, this test is of limited utility because the alternative hypothesis is very restrictive - what if $x_t$ is a stationary **ARMA**$(2, 0, 0)$ or **ARMA**$(1, 0, 1)$ process?

- The **Augmented Dickey-Fuller** test addresses this limitation by testing the null hypothesis that $x_t$ is an **ARIMA** $(p, 1, 0)$ process against the alternative hypothesis that $x_t$ is an **ARMA** $(p + 1, 0, 0)$ process. It assumes that

$$\nabla x_t = a + \kappa x_{t-1} + \phi_1 \nabla x_{t-1} + \cdots + \phi_p \nabla x_{t-p} + w_t, \tag{3}$$

where $w_t \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma_w^2)$. Clearly, if $\kappa = 0$, then $x_t$ is an **ARIMA** $(p, 1, 0)$. If $\kappa \neq 0$, we can rewrite the equation as a stationary **ARIMA** $(p + 1, 0, 0)$ process,

$$x_t = a + (1 + \kappa + \phi_1) x_{t-1} + (\phi_2 - \phi_1) x_{t-2} + \cdots + (\phi_p - \phi_{p-1}) x_{t-p} - \phi_p x_{t-p-1} + w_t.$$

The order $p$ is usually chosen by using AIC, AICc, or SIC/BIC to choose select the "best" **ARIMA** $(p, 1, 0)$ for $x_t$. Once the order $p$ has been chosen, the test statistic is the $t$-statistic $\hat{\kappa}/\text{se}(\hat{\kappa})$ for the the least-squares estimate $\hat{\kappa}$ based on (3), and the approximate distribution of $\hat{\kappa}$ under the null as $n \to \infty$ is still the Dickey-Fuller distribution. This is a much more useful test than the original Dickey-Fuller test, because the alternative hypothesis is the larger class of **ARIMA** $(p + 1, 0, 0)$ models with $p > 0$. Furthermore, because many **ARIMA** $(p + 1, 0, q)$ processes can be well approximated using **ARIMA** $(k, 0, 0)$ processes for some value of $k$, we can think of the alternative hypothesis also approximately including **ARIMA** $(p + 1, 0, q)$ models! This means that the augmented Dickey-Fuller test basically approximately works as a test of the null hypothesis that $x_t$ is an **ARIMA** $(p, 1, q)$ process against the alternative hypothesis that $x_t$ is an **ARMA** $(p + 1, 0, q)$ process.

- The **Phillips-Perron** test is motivated by the concern that although many **ARIMA** $(p + 1, 0, q)$ processes can be well approximated using **ARIMA** $(k, 0, 0)$ processes for some value of $k$, $k$ may need to be too large for this to be practically useful. It tests an even more general null hypothesis against a more general alternative hypothesis, specifically it tests the null hypothesis that $\nabla x_t$ is a stationary process (and $x_t$ is not) against the alternative hypothesis that $x_t$ is a stationary process. We can think of this assuming that

$$x_t - x_{t-1} = a + \kappa x_{t-1} + w_t, \tag{4}$$

where $w_t$ is a stationary process, with a finite number $L$ of nonzero autocorrelations and slightly nonconstant variance. The number $L$ is often chosen in practice as a deterministic function of the length of the time series, $n$. The test statistic is computed from the residuals from the linear regression based on (4), and the approximate distribution of the test statistic under the null as $n \to \infty$ is still the Dickey-Fuller distribution.

The greater generality offered by the Phillips-Perron test does not come for free! Making fewer assumptions about how $x_t$ behaves under the alternative can mean that we may need more data for the test statistic to be approximately Dickey-Fuller distributed under the null, i.e. for the test to actually perform the way we want it to.

These tests are often used to choose the differencing parameter $d$ of **ARIMA**$(p, d, q)$ given an observed time series $\boldsymbol{x}$. Note that it isn't appropriate to choose $d$ using AIC, AICc, SIC/BIC, because differencing reduces the number of available observations, so **ARIMA** models with different values of $d$ are fit using different numbers of observations. We can use one of these non-stationarity tests to select $d$ by specifying a desired level $\alpha$, e.g. $\alpha = 0.05$, and then proceeding as follows:

(i) Conduct a level-$\alpha$ test of non-stationarity of $x_t$.

- Fail to reject non-stationarity $\implies$ Set $k = 1$, proceed to (ii).
- Reject non-stationarity $\implies$ STOP, set $d = 0$.

(ii) Conduct a level-$\alpha$ test of non-stationarity of $\nabla^k x_t$.

- Fail to reject non-stationarity $\implies$ Set $k = k + 1$, proceed to (ii).
- Reject non-stationarity $\implies$ STOP, set $d = k$.

Once an order $d$ has been selected, the same approaches that we used to select $p$ and $q$ for **ARMA**$(p, q)$ models, e.g. minimizing AIC, AICc, SIC/BIC, can be used because we can assume that $\nabla^d x_t$ can be treated as an r **ARMA**$(p, q)$ process.