Econometrics II

Notes - Midterm

William Radaic Peron

EESP-FGV

September 19, 2020

Chapter 1

Introduction

1.1 Motivation

This course will be dedicated to *time series analysis.* Informally, a *time series* is any type of data collected over time – or, more formally, it is the realization of a stochastic process indexed in time. We usually denote the time series as follows:

$$y_1, ..., y_T; \quad \{y_t\}_{i=1}^T; \quad \{y_t\}_t$$

Time series analysis is useful for a number of different applications:

- **Forecasting.**
  - Uni and multivariate models
  - **ARIMA** models: mean and confidence interval forecasting
  - **ARCH** models: variance forecasting – especially useful in finance for volatility and risk
- **Dynamics.** Evaluate the impact of one variable in another over time.
  - Multivariate models including VAR, ECM
  - Contemporaneous lagged structural relations

It is important to address a first and simple question. **Why time series are different from other data?** The answer is also simple but incredibly relevant: *time series observations are not serially independent!*

$$Y_t \not\perp Y_{t-j}$$

In fact, they don't even have to be identically distributed:

$$F_{Y_t} \neq F_{Y_{t-j}}$$

This means that the essential *iid* hypothesis for traditional Econometrics *does not hold.* This means that we'll have to make some adjustments to our methods. That is the task of time series analysis.

1.2 Statistics with dependence

Let's begin with a proper definition of a time series.

1.2.1 Definition of a time series

Suppose that we have a probability space $(\Omega, S, \mathbb{P})$. Ω is the sample space; S is the set of all events; $\mathbb{P}$ is a measure of probability $\mathbb{P} : S \rightarrow [0,1]$. From this, we define a random variable $Y : \Omega \rightarrow \mathbb{R}$. A realization of this r.v. is denoted by $y = Y(\omega)$ with fixed ω.

From this, we can define multiple random variables in the same sample space, indexed by integers:

$$Y = \{..., Y_{t-2}, Y_{t-1}, Y_t, ...\}$$

This is equivalent to writing:

$$Y : \Omega x \mathbb{Z} \rightarrow \mathbb{R}$$

We now arrive at our formal definition of a time series: $\{Y_t, t \in \mathbb{Z}\}$ is a time-indexed stochastic process.

- $Y(\cdot, t) : \Omega \rightarrow \mathbb{R}$ is a r.v. for fixed t.
- $Y(\omega, \cdot) : \mathbb{Z} \rightarrow \mathbb{R}$ is a *sequence of real numbers* for a fixed ω. In other words, this represents the *observed time series.*
- For fixed t, ω, $Y(\omega, t) \in \mathbb{R}$.

1.2.2 Unconditional expectation

An important concept to make clear here is *unconditional expectation.* With fixed t,

$$\mathbb{E}(Y_t) = \int_{-\infty}^{\infty} x f_{Y_t}(x) dx$$

Note the Y_t subscript on the probability density function f_{Y_t}. This means that $\mathbb{E}(Y_t)$ is not calculated with the values assumed by Y_{t-1}, Y_{t+1}. This raises an important problem: *how would you be able to estimate* $\mathbb{E}(Y_t)$*?* Note that we only observe $Y_t = y_t$, i.e., one realization of the r.v.

1.2.3 Statistical dependence

For any random variables X, Y, we can define multiple measures of dependency:

- **Linear:** $Cov(X, Y) \equiv \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$
- **Quadratic:** $Cov(X^2, Y^2)$
- **General:** $Cov(f(X), g(Y))$. This is a measure of covariance between two general functional forms of X and Y.

With this general definition, we arrive at an equivalent definition for independent random variables:

- $F_{X,Y}(x, y) = F_X(x) * F_Y(y)$, i.e., joint pdf is equal to the product of the marginal pdfs.
- $Cov(f(X), g(Y)) = 0$ for every pair of bounded functions f, g.

From this, we now define the *autocovariance and autocorrelation functions.*

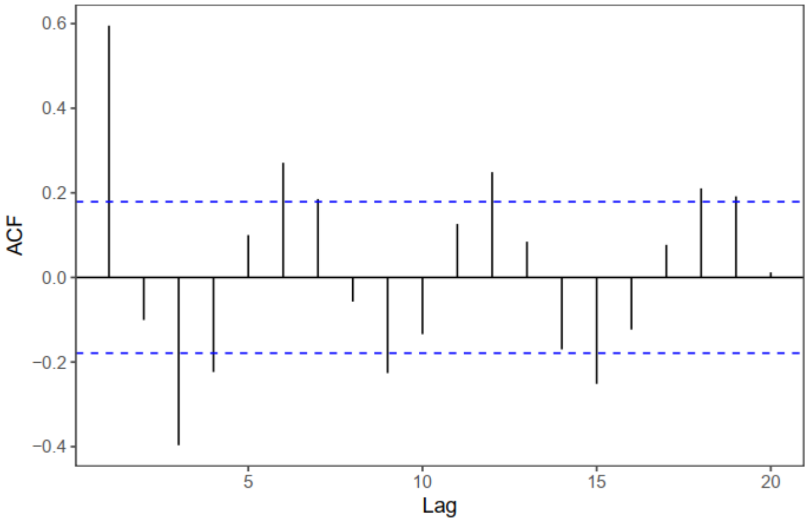
Definition 1.2.1. $\gamma_{j,t} := Cov(Y_t, Y_{t-j})$ *is the* ***autocovariance function*** *for a given time series* $\{Y_t, t \in \mathbb{Z}\}$.

Definition 1.2.2. $\rho_{j,t} := \dfrac{\gamma_{j,t}}{\sqrt{\gamma_{0,t}\gamma_{0,t-j}}}$ *is the* ***autocorrelation function*** *for a given time series* $\{Y_t, t \in \mathbb{Z}\}$.

Note that, if *iid* holds:

$$\gamma_{j,t} = \begin{cases} 0 & j \neq 0, \forall t \\ Var(Y) & otherwise \end{cases}$$

This is an example of an autocorrelation function.



1.3 Asymptotic theory with dependence

Some form of asymptotic theory is needed to enable *any kind of statistical analysis.* Namely, we need to have some form of Law of Large Numbers (LLN) and Central Limit Theorem (CLT) that are analogous to the *iid* environment. This will be achieved in our setting with some conditions called *stationarity* and *ergodicity.*

1.3.1 Stationarity

Definition 1.3.1. *A process* $\{Y_t, t \in \mathbb{Z}\}$ *is* ***strictly stationary*** *if, for all finite set of indexes* $\{t_1, ..., t_r\}$ *and for all* $m \in \mathbb{Z}$, $F(y_{t_1}, ..., y_{t_r}) = F(y_{t_1+m}, ..., y_{t_r+m})$ *holds, where* $F(y_{t_1}, ..., y_{t_r})$ *is the joint cdf of* $(Y_{t_1}, ..., Y_{t_r})$.

More informally, a given process is called *strictly stationary* if its statistical properties depend only on the *relative position* between observations, and not its *absolute position.*

We'll usually adopt a weaker definition of stationarity for our models. Henceforth, we will refer to stationarity in this sense.

Definition 1.3.2. *A process* $\{Y_t, t \in \mathbb{Z}\}$ *is* ***stationary*** *(or weakly stationary) if there exists* $\mu \in \mathbb{R}$ *and* $\{\gamma_j\}_{j \in \mathbb{N}}$ *such that:*

- $\mathbb{E}(Y_t) = \mu, \quad \forall t$
- $\mathbb{E}[(Y_t - \mu)(Y_{t-j} - \mu)] = \gamma_j, \quad \forall (t, j) \in \mathbb{N}^2$

Note that, from the second condition in the definition, we have $\mathbb{E}(Y_t - \mu)^2 = \gamma_0 \in \mathbb{R}, \forall t \in \mathbb{N}$. In other words, *the unconditional variance of the time series is constant.*

Some important remarks on stationarity:

- Stationarity does not imply strict stationarity
- Stricy stationarity does not imply stationarity
- Every strictly stationary process with finite variance is stationary
- Every iid process is strictly stationary
- Every strictly stationary process is identically distributed
- A stationary process is not necessarily identically distributed

1.3.2 Ergodicity

Stationarity is not enough to guarantee that we have even a Law of Large Numbers. To see why that is the case, consider the following example:

$$Y_t = X + \varepsilon_t, \quad \varepsilon \sim \mathcal{N}(0, \sigma^2), \quad X \sim \mathcal{N}(0,1), \quad X \perp\!\!\!\perp \varepsilon_t$$

Is this process stationary? No, because the sample *time average* $\bar{y} = \frac{1}{T}\sum_{t=1}^{T} y_t$ does not converge to the population *ensemble average* $\mathbb{E}(Y_t) = \mu$.

We need some condition that guarantees that the dependence structure of the time series decays as the observation get further from each other. That is the intuition behind *ergodicity.*

Definition 1.3.3. *A strictly stationary process $\{Y_t, t \in \mathbb{Z}\}$ is called* **ergodic** *if*

$$\lim_{J\to\infty} \frac{1}{J} \sum_{j=1}^{J} Cov[f(X_1), g(X, j)] = 0,$$

for all pairs of bounded functions f, g.

This is a kind of mean asymptotic independence, in which the asymptotic independence would be defined by $Cov[f(X_1), g(X_J)] \to 0$ as $J \to \infty$.

Now, we can define a Law of Large Numbers – also called the *Ergodic Theorem.*

Theorem 1.3.1. *Given an ergodic stochastic process $\{Y_t, t \in \mathbb{Z}\}$ such that $\mathbb{E}|Y_1| < \infty$,*

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} Y_t = \mathbb{E}(Y_1) \quad almost\, sure$$

This theorem is the generalization of the strong LLN. However, it presupposes *strict stationarity,* which is a very strong assumption most of the time. Fortunately, this theorem gave rise to other definitions that arrive at our objective, namely, a LLN for the first two moments.

Definition 1.3.4. *A stationary process $\{Y_t, t \in \mathbb{Z}\}$ is said to be* **ergodic for the mean** *if*

$$\frac{1}{T} \sum_{t=1}^{T} Y_t \to_p \mathbb{E}(Y_t), \quad T \to \infty$$

Definition 1.3.5. *A stationary process $\{Y_t, t \in \mathbb{Z}\}$ is said to be* ***ergodic for the second moment*** *if, for every j,*

$$\frac{1}{T-j}\sum_{t=j+1}^{T} Y_t Y_{t-j} \to_p \mathbb{E}(Y_t), \quad T \to \infty$$

Proposition 1.3.2. $\sum_{j=0}^{\infty} |\gamma_j| < \infty$ *is a suficient condition for ergodicity for the mean.*

Proof. Let $Z_t := Y_t - \mu$ and $\bar{Z}_t := \frac{1}{T}\sum_{t=1}^{T} Z_t$, where $\{Y_t, t \in \mathbb{Z}\}$ is a stationary process. We will show that $\bar{Z}_t$ converges to 0 in mean square.

$$\begin{aligned}
\mathbb{E}\left(\bar{Z}_T^2\right) &= \mathbb{E}\left[\left(\frac{1}{T}\sum_{t=1}^{T} Z_t\right)\left(\frac{1}{T}\sum_{t=1}^{T} Z_t\right)\right] = \frac{1}{T^2}\mathbb{E}\left(\sum_{s=1}^{T}\sum_{t=1}^{T} Z_s Z_t\right) \\
&= \frac{1}{T^2}\sum_{s=1}^{T}\sum_{t=1}^{T}\mathbb{E}\left(Z_s Z_t\right) = \frac{1}{T^2}\sum_{s=1}^{T}\sum_{t=1}^{T}\gamma_{s-t} = \frac{1}{T}\sum_{j=-T+1}^{T-1}\frac{T-|j|}{T}\gamma_j \\
&\leq \frac{1}{T}\sum_{j=-T+1}^{T-1}\frac{T-|j|}{T}|\gamma_j| \leq \frac{1}{T}\sum_{j=-T+1}^{T-1}|\gamma_j| \to 0
\end{aligned}$$

□

1.3.3 A Central Limit Theorem for time series

The conditions that guarantee the existence of a CLT for stationary and ergodic processes are much more envolving than in the *iid* environment. However, we have a relatively simple result that will be useful to us in time series analysis. It will now be presented without proof.

Theorem 1.3.3. *Let $\{Y_t, t \in \mathbb{Z}\}$ be a* ***linear*** *stationary process, i.e., that can be written in the form $Y_t = \mu + \sum_{j=0}^{\infty}\psi_j \epsilon_{t-j}$, where $\varepsilon \sim_{iid} (0, \sigma^2)$ and $\sum_{j=-\infty}^{\infty}|\psi_j| < \infty$. Then,*

$$\sqrt{T}(\bar{Y}_t - \mu) \to_d \mathcal{N}(0, \omega^2),$$

where $\omega^2 := \sum_{j=-\infty}^{\infty}\gamma_j < \infty$

Chapter 2

ARMA Models

ARMA is a class of models that we'll employ frequently in time series analysis. Let's begin with some definitions.

2.1 White noise

We call *white noise* stationary time series with mean zero that do not have serial correlation.

Definition 2.1.1. *$\{Y_t, t \in \mathbb{Z}\}$ is **white noise**, denoted by $Y_t \sim wn(0, \sigma^2)$, if*

$$\mathbb{E}(Y_t) = 0; \quad \mathbb{E}(Y_t, Y_{t-j}) = \begin{cases} \sigma^2 & j = 0 \\ 0 & j \neq 0 \end{cases}$$

This is the most simple time series – except for the *iid* case, where independence also holds. It will be the building block for a number of processes that we will study.

2.2 Moving Average processes

Let's begin with the simplest form of MA processes: MA(1).

Definition 2.2.1. *A stationary process $\{Y_t, t \in \mathbb{Z}\}$ is called **MA(1)**, or a **moving average of order 1**, if it follows the following form:*

$$Y_t = c + \varepsilon_t + \theta\varepsilon_{t-1}, \varepsilon_t \sim wn(0, \sigma^2)$$

2.2.1 Moments of an MA(1) model