Recitation One- Time Series

Nathaniel Mark

February 4, 2019

My recitations will have a dual mandate:
1) To solidify understanding of the lectures through examples, sample problems, and more in-depth overviews of specific topics.
2) To provide necessary background skills (both analytical and coding-related) to complete your problem sets.

1 What is the point of time series?

We want to do two things with times series:
1) Figure out how the true world behaves.
e.g. How does putting a subway line into a neighborhood effect the neighborhood's development over time?
2) Predict what the world will look like in the future.
e.g. what will housing prices in a neighborhood look be in 5 years?

We do this by **estimating a statistical model**: from a set of possible probability distributions, we find the one that looks like it generates our **data**. We can then use that estimated model to answer questions about the world and forecast.

2 Statistical Models: Definitions and Examples

A **Statistical Model** is a set of (possibly joint) probability distributions. We denote this as

$$\{P_\theta\}_{\theta \in \Theta}$$

where θ is called the **parameter**. Each probability distribution in the statistical model is represented (indexed) by a value of θ. The set of possible values of θ is written as Θ. Θ is called the **parameter space**. We assume that the true data we observe is drawn from one of probability distributions in the statistical model.

Under the assumption that our statistical model is correct, note that Θ is the set of parameter values that are possible and $\{P_\theta\}_{\theta\in\Theta}$ is the set of probability distributions that possible. Every individual element of Θ corresponds exactly to one element of $\{P_\theta\}_{\theta\in\Theta}$. Therefore, the number of elements in Θ is the same as the number of elements in $\{P_\theta\}_{\theta\in\Theta}$.

Example One:

Set up: We have a box of oranges, but do not know if they are from California or from Florida. We know oranges from Florida have a weight distributed N(3,1) and oranges from California have weight distributed N(2,1). What would be an appropriate statistical model of this situation?

Statistical Model: {N(2,1) and N(3,1)}
Equivalently, this could be written as: $\{N(\mu, 1) and N(\mu, 1)\}_{\mu\in\{2,3\}}$
Parameter: $\theta=\mu$ (mean of the normal distribution)
Parameter Space: {2 and 3}

Example Two:

Set up: A naive equity trader (pre-2008) believes that the S&P increases according to an iid normal distribution, but they do not know the mean or the variance of that normal distribution. The statistical model they would use is:

Statistical Model: $\{N(\mu, \sigma^2)\}_{\mu \text{ is any real number}, \sigma \text{ is any positive real number}}$
Parameter: $\theta=[\mu, \sigma]$ (mean and std dist of the normal distribution)
Parameter Space: $\mathbb{R}\times\mathbb{R}^+$ (which just means μ can be any number and σ can be any positive number)

Example Three:

Set up: An intro to econometrics student thinks the five OLS assumptions apply when discussing the effect of X_i on Y_i. The statistical model they would use is:

Statistical Model: The probability distribution of Y_i is a member of

$$\{N(\beta_0 + \beta_1 X_i, \sigma^2)\}_{\beta_0 \text{ is any real number}, \beta_1 \text{ is any real number}, \sigma \text{ is any positive real number}}$$

Parameter: $\theta=[\beta_0, \beta_1, \sigma]$

Parameter Space: $\mathbb{R} \times \mathbb{R} \times \mathbb{R}^+$ (which just means β_0 can be any number and β_1 can be any number andσ can be any positive number)

Example Four:

We assume that housing prices increase at a steady rate with shocks in each period due to demand, but also are affected by transportation.

Statistical Model:
$P_t = \mu + \phi_1 P_{t-1} + \beta Subway_t + \epsilon_t$,
$\epsilon_t \sim \sigma^2$.
Parameter: $\theta = [\mu, \phi, \beta, \sigma]$
Parameter Space: $\Theta = \mathbb{R}^3 \times \mathbb{R}^+$

3 Time Series Preliminaries

Definition: A random variable ϵ_t is **white noise** if it has the following properties:
1) $E[\epsilon_t] = 0$
2) $Cov(\epsilon_t, \epsilon_t) = Var(\epsilon_t) = \sigma^2 > 0$
3) $Cov(\epsilon_t, \epsilon_s) = 0$ where $t \neq s$

Definition: A random variable ϵ_t is **Gaussian white noise** if it has the following properties:
$\epsilon_t \sim N(0, \sigma^2) i.i.d.$

Definition: A Time Series $\{X_t\}$ is **weakly stationary** if:
1) $E[X_t] = \mu$ (a constant) for all t
2) $Cov(X_t, X_{t-j}) = \gamma(j)$ (does not depend on t) for all $j \geq 0$

Definition: A Time Series $\{X_t\}$ is **strongly stationary** if:
the joint distribution of any sequence $..., X_{t-j}, X_t, X_{t+k}, ...$ does not depend on t.

Why do we care about stationarity? In words, we need to assume that the future looks like the past to be able to use the past to talk about the future. We will go over this in math language later in the course.

4 General MA Model

A general ***q****th order Moving Average Model* (MA(q)) takes the following form:

$$X_t = \mu + \theta_0\epsilon_t + \theta_1\epsilon_{t-1} + \theta_2\epsilon_{t-2} + ... + \theta_q\epsilon_{t-q}$$

where ϵ_t are white noise. Properties of an MA(q):
1) Stable Mean
2) Always Weakly Stationary
3) Autocovariance function of order q+1 or larger is equal to zero. That is, $\gamma_X(h) = 0$ for all $h > q$

I wont write down the mean and autocovariance formulas exactly, because that would give away the homework!

5 General AR Model

A general ***p****th order AutoRegressive Model* (AR(p)) takes the following form:

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + ... + \phi_p X_{t-p} + \epsilon$$

where ϵ_t are white noise. Properties of an AR(p):
1) Stable Mean
2) Weakly Stationary *only under certain conditions*
3) If stationary, autocovariance function tends toward zero as the lag gets larger, but never hits 0.
$\gamma_X(h) \to 0$ as $h \to \infty$

6 Attaining an ACF for an MA

For each lag h,
Step One: Plug in definition of X_t and X_{t-h} into the covariance function.
Step Two: Simplify expression to covariances of ϵs.
Step Three: Apply the white noise definition.

Example: $X_t = \epsilon_t + \frac{1}{2}\epsilon_{t-1}$

$$\gamma_X(0) = Cov(X_t, X_t)$$

Step One:

$$= Cov(\epsilon_t + \frac{1}{2}\epsilon_{t-1}, \epsilon_t + \frac{1}{2}\epsilon_{t-1})$$

Step Two:

$$= Cov(\epsilon_t, \epsilon_t) + \frac{1}{2}Cov(\epsilon_t, \epsilon_{t-1}) + \frac{1}{2}Cov(\epsilon_{t-1}, \epsilon_t) + (\frac{1}{2})^2 Cov(\epsilon_{t-1}, \epsilon_{t-1})$$

Step Three:

$$= \sigma^2 + \frac{1}{2}0 + \frac{1}{2}0, \epsilon_t) + (\frac{1}{2})^2\sigma^2$$

$$= 1.25\sigma^2$$

AND

$$\gamma_X(1) = Cov(X_t, X_{t+1})$$

Step One:

$$= Cov(\epsilon_t + \frac{1}{2}\epsilon_{t-1}, \epsilon_{t+1} + \frac{1}{2}\epsilon_t)$$

Step Two:

$$= Cov(\epsilon_t, \epsilon_{t+1}) + \frac{1}{2}Cov(\epsilon_t, \epsilon_t) + \frac{1}{2}Cov(\epsilon_{t-1}, \epsilon_{t+1}) + (\frac{1}{2})^2 Cov(\epsilon_{t-1}, \epsilon_t)$$

Step Three:

$$= 0 + \frac{1}{2}\sigma^2 + \frac{1}{2}0 + (\frac{1}{2})^2 0$$

$$= \frac{1}{2}\sigma^2$$

$$\gamma_X(h) = Cov(X_t, X_{t+h}) = 0$$

if $h \geq 2$