

# Time Series Cheat Sheet

By Marcelo Moreno - Universidad Rey Juan Carlos  
The Econometrics Cheat Sheet Project

## Basic concepts

### Definitions

**Time series** - is a succession of quantitative observations of a phenomena ordered in time.

There are some variations of time series:

- **Panel data** - consist of a time series for each observation of a cross section.
- **Pooled cross sections** - combines cross sections from different time periods.

**Stochastic process** - sequence of random variables that are indexed in time.

### Components of a time series

- **Trend** - is the long-term general movement of a series.
- **Seasonal variations** - are periodic oscillations that are produced in a period equal or inferior to a year, and can be easily identified on different years (usually are the result of climatology reasons).
- **Cyclical variations** - are periodic oscillations that are produced in a period greater than a year (are the result of the economic cycle).
- **Residual variations** - are movements that do not follow a recognizable periodic oscillation (are the result of eventual phenomena).

### Type of time series models

- **Static models** - the relation between  $y$  and  $x$ 's is contemporary. Conceptually:

$$y_t = \beta_0 + \beta_1 x_t + u_t$$

- **Distributed-lag models** - the relation between  $y$  and  $x$ 's is not contemporary. Conceptually:

$$y_t = \beta_0 + \beta_1 x_t + \beta_2 x_{t-1} + \dots + \beta_s x_{t-(s-1)} + u_t$$

The long term cumulative effect in  $y$  when  $\Delta x$  is:

$$\beta_1 + \beta_2 + \dots + \beta_s$$

- **Dynamic models** - lags of the dependent variable (endogeneity). Conceptually:

$$y_t = \beta_0 + \beta_1 y_{t-1} + \dots + \beta_s y_{t-s} + u_t$$

- Combinations of the above, like the rational distributed-lag models (distributed-lag + dynamic).

## Assumptions and properties

### OLS model assumptions under time series

Under this assumptions, the OLS estimator will present good properties. **Gauss-Markov assumptions** extended applied to time series:

- t1. **Parameters linearity and weak dependence.**
  - a.  $y_t$  must be a linear function of the  $\beta$ 's.
  - b. The stochastic  $\{(x_t, y_t) : t = 1, 2, \dots, T\}$  is stationary and weakly dependent.
- t2. **No perfect collinearity.**
  - There are no independent variables that are constant:  $\text{Var}(x_j) \neq 0, \forall j = 1, \dots, k$
  - There is not an exact linear relation between independent variables.
- t3. **Conditional mean zero and correlation zero.**
  - a. There are no systematic errors:  $E(u | x_1, \dots, x_k) = E(u) = 0 \rightarrow$  **strong exogeneity** (a implies b).
  - b. There are no relevant variables left out of the model:  $\text{Cov}(x_j, u) = 0, \forall j = 1, \dots, k \rightarrow$  **weak exogeneity**.
- t4. **Homoscedasticity.** The variability of the residuals is the same for any  $x$ :  $\text{Var}(u | x_1, \dots, x_k) = \sigma_u^2$
- t5. **No autocorrelation.** Residuals do not contain information about any other residuals:  $\text{Corr}(u_t, u_s | x_1, \dots, x_k) = 0, \forall t \neq s$
- t6. **Normality.** Residuals are independent and identically distributed (**i.i.d.** so on):  $u \sim \mathcal{N}(0, \sigma_u^2)$
- t7. **Data size.** The number of observations available must be greater than  $(k + 1)$  parameters to estimate. (It is already satisfied under asymptotic situations)

### Asymptotic properties of OLS

Under the econometric model assumptions and the Central Limit Theorem:

- Hold t1 to t3a: OLS is **unbiased**.  $E(\hat{\beta}_j) = \beta_j$
- Hold t1 to t3: OLS is **consistent**.  $\text{plim}(\hat{\beta}_j) = \beta_j$  (to t3b left out t3a, weak exogeneity, biased but consistent)
- Hold t1 to t5: **asymptotic normality** of OLS (then, t6 is necessarily satisfied):  $u \sim_a \mathcal{N}(0, \sigma_u^2)$
- Hold t1 to t5: **unbiased estimate** of  $\sigma_u^2$ .  $E(\hat{\sigma}_u^2) = \sigma_u^2$
- Hold t1 to t5: OLS is **BLUE** (Best Linear Unbiased Estimator) or **efficient**.
- Hold t1 to t6: hypothesis testing and confidence intervals can be done reliably.

## Trends and seasonality

**Spurious regression** - is when the relation between  $y$  and  $x$  is due to factors that affect  $y$  and have correlation with  $x$ ,  $\text{Corr}(x_j, u) \neq 0$ . Is the **non-fulfillment of t3**.

### Trends

Two time series can have the same (or contrary) trend, that should lend to a high level of correlation. This, can provoke a false appearance of causality, the problem is **spurious regression**. Given the model:

$$y_t = \beta_0 + \beta_1 x_t + u_t$$

where:

$$y_t = \alpha_0 + \alpha_1 \text{Trend} + v_t$$

$$x_t = \gamma_0 + \gamma_1 \text{Trend} + v_t$$

Adding a trend to the model can solve the problem:

$$y_t = \beta_0 + \beta_1 x_t + \beta_2 \text{Trend} + u_t$$

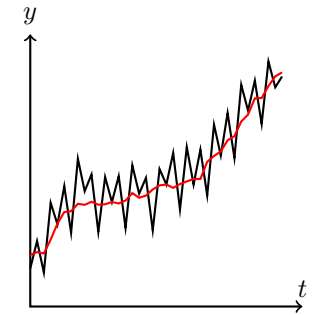
The trend can be linear or non-linear (quadratic, cubic, exponential, etc.)

Another way is make use of the **Hodrick-Prescott filter** to extract the trend and the cyclical component.

### Seasonality

A time series with can manifest seasonality. That is, the series is subject to a seasonal variations or pattern, usually related to climatology conditions.

For example, GDP (black) is usually higher in summer and lower in winter. Seasonally adjusted series (red) for comparison.



- This problem is **spurious regression**. A seasonal adjustment can solve it.

A simple **seasonal adjustment** could be creating stationary binary variables and adding them to the model. For example, for quarterly series ( $Qq_t$  are binary variables):

$$y_t = \beta_0 + \beta_1 Q2_t + \beta_2 Q3_t + \beta_3 Q4_t + \beta_4 x_{1t} + \dots + \beta_k x_{kt} + u_t$$

Another way is to seasonally adjust (sa) the variables, and then, do the regression with the adjusted variables:

$$z_t = \beta_0 + \beta_1 Q2_t + \beta_2 Q3_t + \beta_3 Q4_t + v_t \rightarrow \hat{v}_t + E(z_t) = \hat{z}_t^{sa}$$

$$\hat{y}_t^{sa} = \beta_0 + \beta_1 \hat{x}_{1t}^{sa} + \dots + \beta_k \hat{x}_{kt}^{sa} + u_t$$

There are much better and complex methods to seasonally adjust a time series, like the **X-13ARIMA-SEATS**.

## Autocorrelation

The residual of any observation,  $u_t$ , is correlated with the residual of any other observation. The observations are not independent. Is the **non-fulfillment** of **t5**.

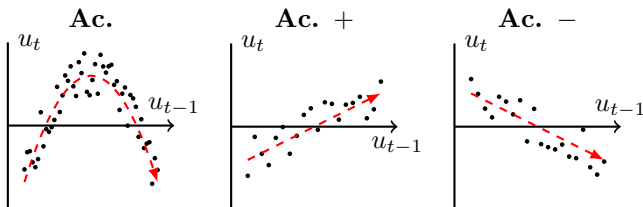
$$\text{Corr}(u_t, u_s \mid x_1, \dots, x_k) = \text{Corr}(u_t, u_s) \neq 0, \forall t \neq s$$

### Consequences

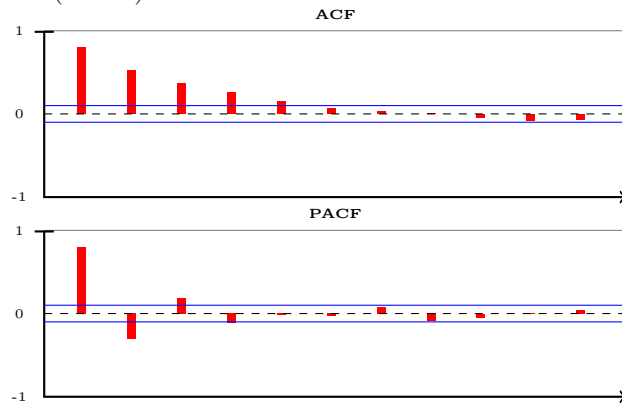
- OLS estimators still are unbiased.
- OLS estimators still are consistent.
- OLS is **not efficient** anymore, but still a LUE (Linear Unbiased Estimator).
- **Variance estimations** of the estimators are **biased**: the construction of confidence intervals and the hypothesis testing is not reliable.

### Detection

- **Scatter plots** - look for scatter patterns on  $u_{t-1}$  vs.  $u_t$ .



- **Correlogram** - composed – Y axis: correlation [-1, 1]. of the autocorrelation func- – X axis: lag number. tion (ACF) and the partial – Blue lines:  $\pm 1.96/T^{0.5}$  ACF (PACF).



Conclusions differ between autocorrelation processes.

- **MA(q) process.** **ACF**: only the first  $q$  coefficients are significant, the remaining are abruptly canceled.
- **AR(p) process.** **ACF**: attenuated exponential fast decay or sine waves. **PACF**: only the first  $p$  coefficients are significant, the remaining are abruptly canceled.
- **ARMA(p, q) process.** **ACF** and **PACF**: the coefficients are not abruptly canceled and presents a fast decay.

If the ACF coefficients do not decay rapidly, there is a clear indicator of lack of stationarity in mean, which would lead to take first differences in the original series.

- **Formal tests** - Generally,  $H_0$ : No autocorrelation.

Supposing that  $u_t$  follows an AR(1) process:

$$u_t = \rho_1 u_{t-1} + \varepsilon_t$$

where  $\varepsilon_t$  is white noise.

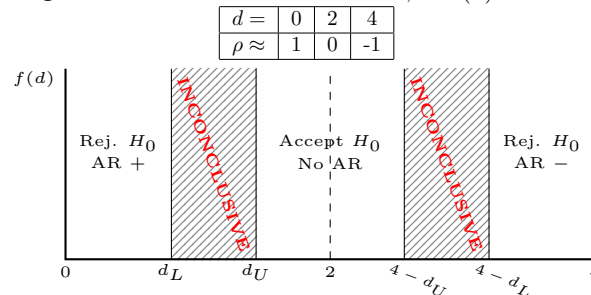
- **AR(1) t test** (exogenous regressors):

$$t = \frac{\hat{\rho}_1}{\text{se}(\hat{\rho}_1)} \sim t_{T-k-1, \alpha/2}$$

- \*  $H_1$ : Autocorrelation of order one, AR(1).
- **Durbin-Watson statistic** (exogenous regressors and residual normality):

$$d = \frac{\sum_{t=2}^n (\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=1}^n \hat{u}_t^2} \approx 2 \cdot (1 - \hat{\rho}_1), 0 \leq d \leq 4$$

- \*  $H_1$ : Autocorrelation of order one, AR(1).



- **Durbin's h** (endogenous regressors):

$$h = \hat{\rho} \cdot \sqrt{\frac{T}{1-T \cdot v}}$$

where  $v$  is the estimated variance of the coefficient associated to the endogenous variable.

- \*  $H_1$ : Autocorrelation of order one, AR(1).
- **Breusch-Godfrey test** (endogenous regressors): it can detect MA( $q$ ) and AR( $p$ ) processes ( $\varepsilon_t$  is w. noise):
- \* MA( $q$ ):  $u_t = \varepsilon_t - m_1 u_{t-1} - \dots - m_q u_{t-q}$
- \* AR( $p$ ):  $u_t = \rho_1 u_{t-1} + \dots + \rho_p u_{t-p} + \varepsilon_t$

Under  $H_0$ : No autocorrelation:

$$T \cdot R_{\hat{u}_t}^2 \underset{a}{\sim} \chi_q^2 \quad \text{or} \quad T \cdot R_{\hat{u}_t}^2 \underset{a}{\sim} \chi_p^2$$

- \*  $H_1$ : Autocorrelation of order  $q$  (or  $p$ ).

- **Ljung-Box Q test**:

- \*  $H_1$ : There is autocorrelation.

### Correction

- Use OLS with a variance-covariance matrix estimator that is **robust to heterocedasticity and autocorrelation** (HAC), for example, the one proposed by **Newey-West**.
- Use **Generalized Least Squares** (GLS). Supposing  $y_t = \beta_0 + \beta_1 x_t + u_t$ , with  $u_t = \rho u_{t-1} + \varepsilon_t$ , where  $|\rho| < 1$  and  $\varepsilon_t$  is white noise.
  - If  $\rho$  is **known**, use a **quasi-differentiated model**:
 
$$y_t - \rho y_{t-1} = \beta_0(1 - \rho) + \beta_1(x_t - \rho x_{t-1}) + u_t - \rho u_{t-1}$$

$$y_t^* = \beta_0^* + \beta_1^* x_t^* + \varepsilon_t$$
 where  $\beta_1' = \beta_1$ ; and estimate it by OLS.
  - If  $\rho$  is **not known**, estimate it by -for example the **Cochrane-Orcutt iterative method** (Prais-Winsten method is also good):
    1. Obtain  $\hat{u}_t$  from the original model.
    2. Estimate  $\hat{u}_t = \rho \hat{u}_{t-1} + \varepsilon_t$  and obtain  $\hat{\rho}$ .
    3. Create a quasi-differentiated model:
 
$$y_t - \hat{\rho} y_{t-1} = \beta_0(1 - \hat{\rho}) + \beta_1(x_t - \hat{\rho} x_{t-1}) + u_t - \hat{\rho} u_{t-1}$$

$$y_t^* = \beta_0^* + \beta_1^* x_t^* + \varepsilon_t$$
 where  $\beta_1' = \beta_1$ ; and estimate it by OLS.
    4. Obtain  $\hat{u}_t^* = y_t - (\hat{\beta}_0^* + \hat{\beta}_1^* x_t) \neq y_t - (\hat{\beta}_0^* + \hat{\beta}_1^* x_t^*)$ .
    5. Repeat from step 2. The algorithm ends when the estimated parameters vary very little between iterations.
- If not solved, look for **high dependence** in the series.

## Exponential smoothing

$$f_t = \alpha y_t + (1 - \alpha) f_{t-1}$$

where  $0 < \alpha < 1$  is the smoothing parameter.

## Predictions

Two types of predictions:

- Of the mean value of  $y$  for a specific value of  $x$ .
- Of an individual value of  $y$  for a specific value of  $x$ .

## Stationarity

Stationarity allows to correctly identify relations –that stay unchanged with time– between variables.

- **Stationary process** (strict stationarity) - if any collection of random variables is taken and shifted  $h$  periods (time changes), the joint probability distribution should stay unchanged.
- **Non-stationary process** - for example, a series with trend, where at least the mean changes with time.
- **Covariance stationary process** - its a weaker form of stationarity:
  - $E(x_t)$  is constant.      –  $\text{Var}(x_t)$  is constant.
  - For any  $t, h \geq 1$ , the  $\text{Cov}(x_t, x_{t+h})$  depends only of  $h$ , not of  $t$ .

## Weak dependence

Weak dependence replaces the random sampling assumption for time series.

- An stationary process  $\{x_t\}$  is **weakly dependent** when  $x_t$  and  $x_{t+h}$  are almost independent as  $h$  increases without a limit.
- A covariance stationary process is **weakly dependent** if the correlation between  $x_t$  and  $x_{t+h}$  tends to 0 fast enough when  $h \rightarrow \infty$  (they are not asymptotically correlated).

Weakly dependent processes are known as **integrated of order zero**,  $I(0)$ . Some examples:

- **Moving average** -  $\{x_t\}$  is a moving average of order  $q$ ,  $MA(q)$ :

$$x_t = e_t + m_1 e_{t-1} + \dots + m_q e_{t-q}$$

where  $\{e_t : t = 0, 1, \dots, T\}$  is an *i.i.d.* sequence with zero mean and  $\sigma_e^2$  variance.

- **Autoregressive process** -  $\{x_t\}$  is an autoregressive process of order  $p$ ,  $AR(p)$ :

$$x_t = \rho_1 x_{t-1} + \dots + \rho_p x_{t-p} + e_t$$

where  $\{e_t : t = 1, 2, \dots, T\}$  is an *i.i.d.* sequence with zero mean and  $\sigma_e^2$  variance.

**Stability condition:** if  $1 - \rho_1 z - \dots - \rho_p z^p = 0$  for  $|z| > 1$  then  $\{x_t\}$  is an  $AR(p)$  stable process that is weakly dependent. For  $AR(1)$ , the condition is:  $|\rho_1| < 1$ .

- **ARMA process** - is a combination of  $AR(p)$  and  $MA(q)$ ;  $\{x_t\}$  is an  $ARMA(p, q)$ :

$$x_t = e_t + m_1 e_{t-1} + \dots + m_q e_{t-q} + \rho_1 x_{t-1} + \dots + \rho_p x_{t-p}$$

## Unit roots

A process is  $I(d)$ , that is, integrated of order  $d$ , if applying differences  $d$  times makes the process stationary.

When  $d \geq 1$ , the process is called a **unit root process** or it is said to have an unit root.

A process have an unit root when the stability condition is not met (there are roots on the unit circle).

### Strong dependence

Most of the time, economics series are strongly dependent (or high persistent). Some examples of **unit root**  $I(1)$ :

- **Random walk** - an  $AR(1)$  process with  $\rho_1 = 1$ .

$$y_t = y_{t-1} + e_t$$

where  $\{e_t : t = 1, 2, \dots, T\}$  is an *i.i.d.* sequence with zero mean and  $\sigma_e^2$  variance.

- **Random walk with a drift** - an  $AR(1)$  process with  $\rho_1 = 1$  and a constant.

$$y_t = \beta_0 + y_{t-1} + e_t$$

where  $\{e_t : t = 1, 2, \dots, T\}$  is an *i.i.d.* sequence with zero mean and  $\sigma_e^2$  variance.

### Unit root tests

Test	$H_0$	Reject $H_0$
ADF	$I(1)$	$\text{tau} < \text{Critical value}$
KPSS	$I(0)$ level	$\text{mu} > \text{Critical value}$
	$I(0)$ trend	$\text{tau} > \text{Critical value}$
Phillips-Perron	$I(1)$	$Z\text{-tau} < \text{Critical value}$
Zivot-Andrews	$I(1)$	$\text{tau} < \text{Critical value}$

### From unit root to weak dependence

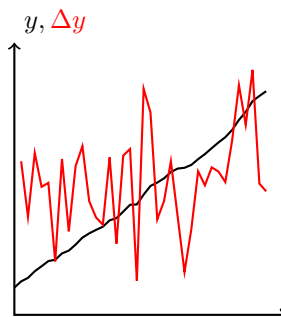
Integrated of **order one**,  $I(1)$ , means that **the first difference** of the process is **weakly dependent** or  $I(0)$  (and usually, stationary). For example, let  $\{y_t\}$  be a random walk:

$$\Delta y_t = y_t - y_{t-1} = e_t$$

where  $\{e_t\} = \{\Delta y_t\}$  is *i.i.d.*

Note:

- The first difference of a series removes its trend.
- Logarithms of a series stabilizes its variance.



### From unit root to percentage change

When an  $I(1)$  series is strictly positive, it is usually converted to logarithms before taking the first difference to obtain the (approx.) percentage change of the series:

$$\Delta \log(y_t) = \log(y_t) - \log(y_{t-1}) \approx \frac{y_t - y_{t-1}}{y_{t-1}}$$

## Cointegration

When **two series are  $I(1)$ , but a linear combination of them is  $I(0)$** . If the case, the regression of one series over the other is not spurious, but expresses something about the long term relation. Variables are called cointegrated if they have a common stochastic trend.

For example,  $\{x_t\}$  and  $\{y_t\}$  are  $I(1)$ , but  $y_t - \beta x_t = u_t$  where  $\{u_t\}$  is  $I(0)$ . ( $\beta$  is the cointegrating parameter).

### Cointegration test

Following the example above:

1. Estimate  $y_t = \alpha + \beta x_t + \varepsilon_t$  and obtain  $\hat{\varepsilon}_t$ .
2. Perform an ADF test on  $\hat{\varepsilon}_t$  with a modified distribution.

The result of this test is equivalent to:

- $H_0: \beta = 0$  (no cointegration)
- $H_1: \beta \neq 0$  (cointegration)

if test statistic  $>$  critical value, reject  $H_0$ .

## Heterocedasticity on time series

The **assumption** affected is **t4**, which leads **OLS to be not efficient**.

Use tests like Breusch-Pagan or White's, where  $H_0$ : No heterocedasticity. It is **important** for the tests to work that there is **no autocorrelation**.

### ARCH

An autoregressive conditional heterocedasticity (ARCH), is a model to analyze a form of dynamic heterocedasticity, where the error variance follows an  $AR(p)$  process.

Given the model:  $y_t = \beta_0 + \beta_1 z_t + u_t$  where, there is  $AR(1)$  and heterocedasticity:

$$E(u_t^2 | u_{t-1}) = \alpha_0 + \alpha_1 u_{t-1}^2$$

### GARCH

A general autoregressive conditional heterocedasticity (GARCH), is a model similar to ARCH, but in this case, the error variance follows an  $ARMA(p, q)$  process.