Teste intercalar - cadeira de Séries Temporais

Parte 1

Pedro Miguel Sousa Magalhães

2021-12-06

Question 1:

Define and explain the importance of the following concepts:

a) stationarity
b) Autocorrelation function
c) Standard error of the estimated coefficients of an AR model applied to a time series

Answer:

a)

A time series is weak stationarity when both conditions are met:

i) The expected value $E(x_t)$ is constant and non dependent on the value of t
ii) The autocovariance between two points of the time series in time (x_t, x_s) depends only on the difference time between then $(t - s)$

Statonairty means that statistical properties of a time series do not change over time. On a non stationairty time series, measures like expected value and forecast depends on the time interval used.

b)

The Auto Correlation Function (ACF) measures the linear predictability of a series at time t given values at time s. It measures how past observations explain current observations following a linear relation between them. It is formally defined as follows:

$$\rho(s,t) = \frac{\gamma(s,t)}{\sqrt{\gamma(s,s)\gamma(t,t)}}$$

Therefore if

$$\rho(s,t) = 1$$

then the relationship between x_t and $x - S$ can be *perfectly* explained by $x_t = \beta_0 + \beta_1 x_t$.

c)

The standard error of the coefficient measures how precisely the model estimates the coefficient's value. On a time series it is commom for errors/residuals to also follow a time series structure (hence erros are not iid). If the time structure of errors are not taken into considerations then the coefficients can be wrongly estimated.

Question 2:

Consider the monthly time series X_t satisfying:

$$X_t = 2 + 0,5t + S_t + e_t$$

. With e_t equals White Noise with $\sigma^2 = 1,5$

a) Show that X_t is not stationary
b) Show that $Y_t = (1 - B^4)X_t$ is stationary

Answer:

a)

Given the definition presented for weak stationarity on question 1a) above we test if both condition apply:

$$E(X_t) = E(2) + E(0,5t) + E(S_t) + E(e_t) \Leftrightarrow E(X_t) = 2 + 0,5t + S_t$$

Therefore the expected value is neither constant or independent of time and therefore, since it violates one of the two necessary conditions, X_t is not stationary.

b)

Based on the concept of backshift operator we can rewrite the expression as such:

$$Y_t = (1-B^4)X_t = X_t - X_{t-4} = (2+0,5t+S_t+e_t) - (2+0,5(t-4)+S_{t-4}+e_{t-4}) = 2+(S_t-S_{t-4})+(e_t-E_{t-4}) = 2+(e_t-e_{t-4})$$

Since $S_t = S_{t-4}$. Therefore:

$$E(Y_t) = E[2 + (e_t - e_{t-4})] = 2$$

$E(Y_t)$ is constant and independent of t fulfilling the first condition for stationary. Lets verify if second condition stands

$$\gamma_y(t,s) = Cov(y_t, y_s) = E\left[(y_t - \mu_{y_t})(y_s - \mu_{y_s})\right] = E(y_t y_s) - 2E(y_t) - 2E(y_s) + 4 = E(y_t y_s) - 4$$

Since y_t and y_s are not iid as by definition of a time series pattern the solutions varies given the differnce between t and s.

$$\begin{cases} E(y_t^2) & t = s \\ E(y_t y_s) & t \neq s \end{cases}$$

Question 3:

a) he sample acf of a time series X1,. . . ,X100 is represented in the plot below. Is this time series white noise? Justify

Answer:

a)

For *White Noise (WN)* series we expect the autocorrelation to be close to zero for all lags. From the ACF plot LAG 2 surpasses the 0,2 threshold hence we can conclude that at lag 2 the correlation is significant.

b)

As stated on the previous answer given a WN we expect the ACF to be close to zero for all LAG's. To study the relevance of the ACF per lag we use 0.6324555 as threshold and we conclude that, based on the available information, the ACF does not surpass this threshold and it converges toward zero, therefore we can conclude this is a White Noise process.

c)

Given the characteristics of a **AR(1)** like $x_t = \phi x_{t-1} + e_t$ with e_t following a white noise process we know:

$$\rho(1) = \frac{\gamma_x(1)}{\gamma_x(0)} = \phi$$

Given that $\gamma_x(0) = \sigma_e^2 \frac{1}{(1-\phi\text{ł}2)}$. Therefore the coeficients of x_{t-1} for each model will be as follows:

Modelo 1: $a = \rho(1) = 0.8$ Modelo 2: $a = \rho(1) = -0,6$

Question 4:

The plot below represents the p-values of the Ljung-Box test to the residuals obtained after fitting an AR(3) to a time series. Explain why the lag (H) starts at the value 4. Write down the null hypothesis, the test statistic and its distribution.

Answer:

It starts at 4 because the test $LAG - 3$ degrees of freedom.

The *Ljung-Box test* null hypothesis is $H0 : Q \sim \chi^2_{H-p-q}$ and Q_statistic given by $Q = n(n+2)\sum_{h=1}^{H} \frac{\rho_e^2(h)}{n-h}$