ARIMA fitting: Daily female births in California in 1959

PRACTICAL TIME SERIES ANALYSIS

THISTLETON AND SADIGOV

Objectives

- Fit an ARIMA model to a real world data set
- Judge various fitting tools such as ACF, PACF and AIC
- Examine Ljung-Box test for testing autocorrelation in a time series

Modeling

- Trend suggests differencing
- Variation in variance suggests transformation
- Common transformation: log, then differencing
- It is also known as log-return
- ACF suggests order of moving average process (q)
- PACF suggests order of autoregressive process (p)
- Akaike Information Criterion (AIC)
- Sum of squared errors (SSE)
- Ljung-Box Q-statistics
- Estimation!

Daily female births in CA, 1959

- Time Series Data Library (TSDL)
- Created by Rob Hyndman, Professor of Statistics at Monash University, Australia.
- Link: https://datamarket.com/data/list/?q=provider%3Atsdl
- Category: Demography
- Name: Daily total female births in California, 1959
- 01 January 1959 – 31 December 1959
- Daily time series

Obtaining the data

- Click on the link: https://datamarket.com/data/set/235k/daily-total-female-births-in-california-1959#!ds=235k&display=line
- Export as CSV file
- Open the file, clean up the bottom row.
- Put the file into your working directory and read it to R
- OR read it directly from its path to R

ARIMA(0,1,2)

Then we have,

$$(1-B)X_t = 0.015_{0.015} + Z_t - 0.8511_{0.0496}\, Z_{t-1} - 0.1113_{0.0502}\, Z_{t-2}$$

where moving average coefficients are significant in the level of 0.05, and indices are standard errors.

Thus, the fitted model is

$$X_t = X_{t-1} + 0.015 + Z_t - 0.8511\, Z_{t-1} - 0.1113\, Z_{t-2}$$

where

$$Z_t \sim Normal\,(0, 49.08)$$

What We've Learned

- How to fit an ARIMA model to a real world data set using various fitting tools such as ACF, PACF and AIC
- Examine Ljung-Box test for testing correlation in a time series