Name: Your name here
Due: 2024/11/27

Day 22 - Lab: ARIMA models

Introduction

In this assignment, we will get more practice with time series inference using GLS and ARIMA models. This lab is graded on completion - just make sure you turn something in by 11/27.

For this lab, we will focus on two data sets: one to practice forecasting via `auto.arima`, and the other to practice GLS. The data we will use for forecasting concerns the global mean methane abundance for marine surface sites, courtesy of NOAA[1].

---

[1] Link to the NOAA site. You do not need to download anything from this link.

Question 1

Suppose we are interested in determining the last month during which we are 95% confident that the mean atmospheric methane abundance is at most 2000. Fit a model to answer this question, demonstrate that there is no residual autocorrelation from this fit, and produce a figure that showcases your fitted model.

Question 2

Let us return to GLS to get a bit more practice answering research questions. To motivate this exercise, we will use a cool data set containing small rodent abundance and rainfall data in Chile. From a rather interesting paper[2]:

> Rodent outbreaks or irruptions in semiarid Chile are associated with rainfall pulses driven by the El Niño Southern Oscillation (ENSO). During the last decade, north-central Chile has experienced an almost uninterrupted sequence of dry years, the so-called megadrought, which had led to a new ecological situation in this region. We employ a diagnostic approach to analyze the abundance of data regarding two rodent species, Phyllotis darwini and Abrothrix olivacea, using a 33-yr time series spanning from 1987 to 2019. Our population dynamic models provide evidence of competitive interactions within and among both species of rodents. This result is novel since rainfall variability influences the degree of interspecific competition and is asymmetric. The diagnostic approach used here offers a way to develop simple population models that are useful for understanding the causes of population fluctuations and for predicting population changes under a climate change scenario.

Here is a bit of information about each column from the metadata:

- `PD` = Phyllotis darwini minimun number alive
- `PDD` = Phyllotis darwini density (minimun number alive/trapping area)
- `XPD` = natural logarithm of P. darwini density
- `AO` = Abrothrix olivacea minimun number alive
- `AOD` = Abrothrix olivacea density (minimun number alive/trapping area)
- `XAO` = natural logarithm of Abrothrix olivacea density
- `RPD` = logarithmic (per capita) reproductive rate of Phyllotis darwini
- `RAO` = logarithmic (per capita) reproductive rate of Abrothrix olivacea
- `P` = Annual accumulated rainfall (mm)
- `P1` = one year lagged annual accumulated rainfall (mm)

[2] Link to the repository. You do not need to download anything from this website.

Is there evidence that minimum number of rodents alive depends upon the rainfall after accounting for differences between species? **Hint**: use the log transformed minimum number alive.