LTCC: Time Series Analysis
2024–2025
Take-home Exam

Part I: R (50%)

Dataset

Mental health is becoming an increasingly important concern in today's society. In this part, we look at the number of suicides by month in England and Wales from January 1981 to December 2015:

https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/adhocs/006905suicidesbymonthofoccurrence1981to2015

For the following analysis, we denote the series as $\{X_t\}$.

(Hints: data pre-processing is needed here; you can do it either directly in Excel or in R. It is also a good idea to take a look at the data file and check its format, before you start answering the questions.)

Questions

There are ten short questions. All of them will be given equal weight (5%).

1. Plot $\{X_t\}$. How many observations are there in this series? Please briefly comment on the visual appearance of this series.

2. Consider differencing the series. Let $Y_t := X_t - X_{t-1}$. Plot $\{Y_t\}$ and briefly comment on the visual appearance of this series.

3. Apply the augmented Dickey–Fuller test on $\{Y_t\}$. Comment on your findings.

4. Plot the ACF estimates of $\{Y_t\}$. Give the estimated values of the ACF at lags $h = 1, 2, 3$.

5. Suppose that we would like to model $\{Y_t\}$ using a MA(q) process, with the order q selected using BIC. Write down your selected model and the estimated values of the parameters.

6. Discuss the adequacy of the selected MA model in view of the estimated ACF.

7. Suppose that we would like to model $\{Y_t\}$ using an ARMA(p, q) process, with the orders p, q selected using AIC. Write down your selected model and the estimated values of the parameters.

8. Suppose we would like to predict the number of suicides in the first two months of 2016 using the model for $\{Y_t\}$ you selected in Question 7. Write down the predicted values.

9. Discuss from a statistical perspective the general trend of $\{X_t\}$ in view of the models from Questions 6 and 7.

10. State two limitations of modelling this particular series $\{Y_t\}$ directly using ARMA.

Part II: Theory and Methodology (50%)

There are two questions. They will be given equal weight (25%).

1. Consider a causal autoregressive process $\{Y_t\}$ of order 2, where

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \epsilon_t, \qquad \text{with } \epsilon_t \overset{i.i.d.}{\sim} N(0, \sigma^2) \text{ and } \sigma > 0.$$

   (a) Given observations $Y_1, \ldots, Y_n$, carefully explain how you would estimate ϕ_1, ϕ_2 and σ^2 using the Yule–Walker equations. Explicit expressions are required here.

   (b) Prove that the PACF of $\{Y_t\}$ at $h = 4$ is zero.

   (c) By considering the asymptotic distribution of the estimates you derived in the previous part or otherwise, construct a 95% confidence interval for the quantity $\phi_1 + \phi_2$. Hence propose a way to test $\phi_1 = -\phi_2$.

   (d) Now suppose that $\phi_1 = -\phi_2$ and $\phi_1 < 0$. Give the range of ϕ_1 values for which $\{Y_t\}$ is invertible.

2. Consider a variant of ARCH(1) where

$$\begin{aligned} Y_t &= \sigma_t \, \epsilon_t, \qquad t = 1, \ldots, n, \\ \sigma_t &= b_0 + b_1 \, |Y_{t-1}|, \end{aligned}$$

   where $\{\epsilon_t\}$ is a sequence of independent Laplace random variables (with zero-mean and unit-variance), and where b_0 and b_1 are positive constants. Assume that $\{Y_t\}$ is both strongly and weakly stationary.

   (a) Suppose that the values of b_0 and b_1 are known.

      i. Compute $E(Y_t)$ and $E(|Y_t|)$.

      ii. Assume that $EY_t^2 < \infty$. Compute the ACF and PACF of both $\{Y_t\}$ and $\{|Y_t|\}$ for all lags.

   (b) Now suppose that the values of b_0 and b_1 are unknown, but you are given the observed time series $\{Y_t\}_{t=1}^n$ that follows this model. Explain how to estimate b_0 and b_1 using (i) least squares and (ii) maximum likelihood.

   (c) Briefly discuss the potential advantages and disadvantages of this model in comparison to the traditional ARCH(1).