UNIVERSITY OF BRITISH COLUMBIA
Department of Statistics
Stat 443: Time Series and Forecasting
Assignment 4: Analysis in the Frequency Domain

The assignment is due on **Thursday April 6** at **9:00pm**.

- Submit your assignment online in the **pdf format** under module "Assignments".
- Include all steps of your derivations as partial marks will be given.
- Please make sure your submission is clear and neat. If you do not know how to typeset mathematical expressions, write your derivations neatly by hand. It is the student's responsibility that the submitted file is in good order (e.g., not corrupted and is what you intend to submit).
- **Late submission penalty**: 1% per hour or fraction of an hour. (In the event of technical issues with submission, you can email your assignment to the instructor to get a time stamp but submit on canvas as soon as it becomes possible to make it available for grading.)

1. In Assignment 2, you have shown that the autocorrelation function for the AR(2) process

$$X_t = \frac{11}{10}X_{t-1} - \frac{3}{10}X_{t-2} + Z_t, \qquad \{Z_t\}_{t\in\mathbb{Z}} \sim WN(0,\sigma^2)$$

is given by

$$\rho(h) = \frac{45}{13}\left(\frac{3}{5}\right)^{|h|} - \frac{32}{13}\left(\frac{1}{2}\right)^{|h|}, \quad h = 0, \pm1, \pm2 \dots.$$

(a) Derive the normalized spectral density function of $\{X_t\}_{t\in\mathbb{Z}}$.

(b) Write down the power spectral density function of $\{X_t\}_{t\in\mathbb{Z}}$.

(c) Plot the normalized spectral density and comment on its behaviour.

2. Given the spectral density function

$$f(\omega) = \frac{1}{8\pi}\left(1 - \frac{32}{25}\cos(\omega) + \frac{3}{5}\cos(2\omega)\right), \quad \omega \in (0,\pi),$$

compute the autocovariance function $\gamma(k)$ of the underlying stochastic process, where $k \in \mathbb{Z}$.

(a) Compute the autocovariance function $\gamma(k)$ at lag $k = 0$.

(b) Compute the autocovariance function $\gamma(k)$ at lag $k = 1$.

(c) Compute the autocovariance function $\gamma(k)$ at lag $k = 2$.

(d) Compute the autocovariance function $\gamma(k)$ for $k > 2$.

(e) Combine the results from (a) to (d) and find the expression for $\gamma(k)$ for $k \in \mathbb{Z}$.

3. (This question must be completed in R Markdown; display all the R code used to perform your data analysis.)

The data file `pelt.txt` gives the annual sales of lynx pelt by the Hudson Bay Company, between 1857 and 1910.

(a) Read the data into `R`, and coerce the data into a time series object. Scale the data by 1000 so that you have sales in thousands. Plot the resulting time series and its sample acf. Comment on what you observe. (Make sure to properly label the axes and provide titles for the plots.)

(b) Plot the raw periodogram for the series. Comment on what you observe and estimate the wavelength and angular frequency for the dominating frequency.

(c) Build a function in `R` that generates the Fourier frequency ω_p for a given time series and given constant $p \in \{0, 1, ..., N/2\}$. Document the inputs and outputs of this function, so that another person would be able to understand how to use your function. What is the output of your function for $p = 8$?

(d) To determine which Fourier frequencies are significant, suppose we were to fit the linear model

$$Y_t = a_0 + a_p \cos(\omega_p t) + b_p \sin(\omega_p t) + \epsilon_t, \qquad t = 1, \ldots, N,$$

where Y_t is the log transform of the time series and $\omega_p = \frac{2\pi p}{N}$ for each $p = 1, 2, ..., N/2$. Assume that $\epsilon_t \sim N(0, \sigma^2)$ for all t and are independent. On fitting the above model for a given p by least squares, a test of the significance of the contribution of frequency ω_p is a test with null hypothesis

$$H_0 : a_p = b_p = 0,$$

that uses the F-test statistic

$$F_p = \frac{\frac{1}{k-1}\sum_{t=1}^{N}(\hat{y}_{t,p} - \overline{y})^2}{\frac{1}{N-k}\sum_{t=1}^{N}(y_t - \hat{y}_{t,p})^2},$$

where k is the number of estimated coefficients in the linear model, N is the number of observations, $\hat{y}_{t,p} = \hat{a}_0 + \hat{a}_p \cos(\omega_p t) + \hat{b}_p \sin(\omega_p t)$ and $\overline{y} = \frac{1}{N}\sum_{t=1}^{N} y_t$. Asymptotically, $F_p \sim F_{2,N-3}$.

Find all Fourier frequencies which are significant at the 95% confidence level.

Hint: Use function `lm()` to fit the linear model. The output of this function can also be used to extract the value of the F-statistic, or compute it directly.

(e) Give the estimated coefficients for the linear model that results from using all significant frequencies found in part (d).

(f) Plot the log transform of the data and the estimated model's fitted values on the same plot in `R`. Remember to properly label the axes, specify a legend and title for the plot.