Programming assignment No. 2

General hints and rules:

- Although the submission of this assignment is not mandatory, you are advised and encouraged to submit your solution, as you can earn bonus points (to be added to your final exam score).
- The solutions can be submitted until Friday, May 24 2024, 11.30 am to tilmann.haertl@uni-konstanz.de. Late submissions will not be considered.
- Program from scratch and don't use 'packages' or 'functions' by other people if not indicated otherwise. Programming on your own helps you to understand the relevant concepts. You may reuse program code that you have written for the first programming assignment.
- We recommend to use the programming language Python* for the following programming problems. The solutions should be handed in form of a code file. Answers to questions requiring text (e.g. for interpretations) should be given as comments in the code.
- Name your code file with the assignment number and your student number (not your name). Example: For programming assignment 2 (PA2) and if your student number is 007007, call the file PA2_007007.py
- Policy with regard to academic dishonesty: Students who wish to work together on assignment material may do so. Please indicate the student numbers of all fellow students that have worked in your group. **Each student must submit her or his individual code/solution file**.
- Write your code general enough such that it can be easily reused. Document your code.
- Avoid loops as much as possible and use matrix language techniques whenever possible.

1) Write a function that simulates data from an AR(1) process $y_t = \alpha y_{t-1} + \varepsilon_t$, where ε_t is iid $N(0, \sigma^2)$. Use a zero initial value, generate $T + 50$ observations and discard the first 50 observations. Your procedure should take α, σ^2 and T as an input and should return a $T \times 1$ vector of observations on y_t. *You may modify the AR(2) function of Computer Tutorial 1*

2) Generate data from the MA(1) process $y_t = \varepsilon_t + \theta\varepsilon_{t-1}$ using, ε_t iid $N(0, 1)$, $\theta = 0.7$ and $T = 500$. Estimate the autocorrelation ρ_3 and test $H_0 : \rho_3 = 0$ on 5% significance level. Implement two versions of the test: (i) using the standard error $1/\sqrt{T}$, (ii) using a standard error based on Bartlett's formula assuming that the data has been generated by an MA(1) process. Repeat this for $M = 1000$ sets of time series and record how often you reject H_0 in test (i) and (ii), respectively. Report the relative rejection frequency over the M Monte Carlo repetitions and interpret your results. *Note: You may use the MA(1) function of Computer Tutorial 1 or a built-in Python function to generate the data. You may also reuse your function from programming assignment 1 (or a built-in Python function) to estimate ρ_3.*

*A download link for Python can be found on ILIAS. There is also a link to a Python documentation.

3) Simulate $T = 100$ observations from an AR(2) process with $c = 0$, $\alpha_1 = 0.6$, $\alpha_2 = 0.2$ and ε_t iid $N(0, \sigma^2)$, $\sigma = 1$. Estimate α_1 and α_2 using the Yule-Walker estimator. Explain your findings. Repeat the exercise with $T = 500$ and compare your outcome with the previous results. *You may use the AR(2) function of Computer Tutorial 1 or a suitable Python package to generate the data.*

4) Consider the following MA(2) process

$$y_t = 1 - 0.5\varepsilon_{t-1} + 0.3\varepsilon_{t-2} + \varepsilon_t.$$

Is the moving average process invertible? Explain. Hint: Use Python to compute the roots of the relevant polynomial.

5) In this exercise you investigate the properties of the LS estimator $\hat{\alpha}$ in a stationary zero mean AR(1) process using a Monte Carlo simulation.

   a) Generate $M = 2000$ sets of time series with $T = 200$ from an AR(1) with $\alpha = 0.4$ and $\sigma^2 = 1$ using your procedure from 1).

   b) For each simulated time series, compute and store the LS estimate $\hat{\alpha}$. Also compute $1/\sqrt{T}\sum_{t=1}^{T} y_{t-1}\varepsilon_t$ and $\sqrt{T}(\hat{\alpha} - \alpha)$.

   c) For each estimate of α record the bias $\hat{\alpha} - \alpha$. Compute the mean bias over the $M = 2000$ replications and provide plots of the sampling distribution of $\hat{\alpha}$.

   d) It can be shown that in a stationary process $1/\sqrt{T}\sum_{t=1}^{T} y_{t-1}\varepsilon_t$ converges to a normal quantity with distribution $N(0, \sigma^4/(1-\alpha^2))$. Moreover, $\sqrt{T}(\hat{\alpha}-\alpha)$ converges to a normal quantity with $N(0, 1-\alpha^2)$. Compare the sampling distributions of $1/\sqrt{T}\sum_{t=1}^{T} y_{t-1}\varepsilon_t$ and $\sqrt{T}(\hat{\alpha}-\alpha)$ from b) with the respective normal distributions by plotting histograms together with the respective normal densities.

   e) Repeat the simulation experiment for $\alpha = 0.9$ and $\alpha = 1$. Compare the mean bias and sampling distributions of $\hat{\alpha}$ with the respective quantities for $\alpha = 0.4$. Also compare the histograms of $\sqrt{T}(\hat{\alpha} - \alpha)$ for different αs. What do you observe?