MCP261 IE Lab I: Jan 17, 2024

Due date: 11:59 PM, Jan 21, 2024

Exercise 2: Discrete-time Markov Chains and Monte Carlo Simulation

Submission guidelines

Each Python file *must* be easy to read: all variables and key quantities of interest must be clearly named, all computations must be commented, all assumptions must be very clearly stated, and finally, you may even add a short explanation of your approach to the problem at the beginning of the .py file. The exercise must be submitted as a zipped folder containing two Python files (.py scripts only – no .txt files or Jupyter notebooks will be entertained) and the Excel file from which you are reading the data. Naming convention for the zipped folder: Name_EntryNum_Exercise#.zip. Naming convention for the Python file for each question: Name_Exercise#_Question#.py Your submission will be heavily penalized if it does not follow the above requirements. Note that the Excel file containing the transition probability matrix must also be provided in the zipped folder so that your code can be run directly from the folder to which your submission is extracted.

Exercise

Consider the Markov chain with the states in the transition probability matrix in the accompanying Excel file. The model represents the progression of patients with schizophrenia who start treatment with a long-acting injectable (LAI) antipsychotic drug. A brief description of patient movement through the model is also given in the file.

Use only the NumPy, SciPy, Pandas and Matplotlib packages for this exercise.

1. (5 marks) Write Python code to generate the average time spent by a patient in each state over a period of 5 years. Also use the costs associated with being in each state given in the file to generate the average total costs spent by a patient over 5 years. Use a time interval of 1 month, and for the initial distribution of patients across states, assume all are in the first state. Assume that transitions take place at the beginning of each month.

2. (5 marks) Conduct a Monte Carlo simulation (1000 iterations) to estimate the mean and standard deviation of the total costs spent by a patient over 5 years. Assume a Gaussian distribution for the total costs associated with each state (with μ = default value of total cost and σ = 10% of the default value of total cost). Use a seed with value 1234 for the random number generator. Report the mean and standard deviation of the total costs spent by a patient over 5 years. Plot the (well-labeled) histogram of costs obtained from the previous question (i.e., using the vector of 1000 cost values from Monte Carlo simulation) using the *matplotlib* package.