

GX5004: HW 4

R and Python have extensive libraries online that can guide you on this assignment. Please feel free to work collectively on the challenging questions. [Total Points: 12]

1. In this exercise, you will work with Markov transition matrices to iterate a system forward through time. [3points]
 - a. Using the rat example from class, together with R or Python, iterate the associated Markov matrix to obtain the transition probabilities two steps ahead, five steps ahead, 10 steps ahead, and 25 steps ahead. (This will recreate the matrices presented in class.)
 - b. Systems can contain so-called “absorbing states.” An absorbing state is a state (or room, to use the classroom example) from which there can be no further transition. In other words, once the state (or room) is entered, the probability of exit is 0. Consider a three-state system with the following transition probabilities.

1	0	0
0.25	0.50	0.25
0	0	1

This system is consistent with two absorbing states: Room A and Room C. Using R or Python, iterate this Markov matrix to obtain the transition probabilities two steps ahead, five steps ahead, 10 steps ahead, and 25 steps ahead. In the limit, what do you think the Markov transition matrix looks like? Continue to iterate the system until the Markov transition matrix converges. Suppose you start the rat in Room B, in the limit, what can you say probabilistically about the evolution of the system?

- c. Challenging question: Consider the maze presented in class, and add two rooms to the right of Room C, labeling them Room D and Room E. In this situation, treat Room A as an absorbing state. If the rat is in any room other than Rooms A or E, it has probability 0.5 of remaining in that room, probability 0.25 of moving left and probability 0.25 of moving right. For Room E, assume probability 0.5 of remaining in that room, and probability 0.5 of moving left. Write out the matrix of Markov transition probabilities. Iterate this matrix forward as many times as is necessary for you to determine empirically its limit. Based on this limit, what can you say about the evolution of the system if the rat begins in Room C? Is there a general conclusion you can draw?
2. Download a revised version of the National Longitudinal Survey of Women (NLSW) data presented in class, called “union_pred.dta”, from the course website. These data are formatted as a Stata dataset. [6points]
 - a. Read this dataset into R or Python. You will see that it is smaller than the earlier version of the dataset but has an additional variable called “prior_union”. For each individual in the sample, this variable is prior union status, for which a 0 indicates “not in union” and a 1 indicates “in union”. Using R or Python, create a two-by-two table that relates prior union status to current union status. Your results should match what was presented in class. Again using R or Python, create a two-by-two matrix of Markov transition probabilities based on these results (either in percentage or decimal format).

- b. Using R or Python, estimate the logit model presented in class, using prior union status as a characteristic. Present the results in a “nice” table.
 - c. Challenging questions (c through f): Now we will simulate transition probabilities using your logit model. Using R or Python, reset your “prior_union” variable so that it takes on value 0 for all observations. For each observation, generate the predicted probability (or predicted value of the logit model) that someone is in a union and call it “pred01”. This prediction is the probability someone joins a union conditional on not having been in a union in the prior time period. Generate a variable called pred00 that is equal to $1 - \text{pred01}$. This prediction is the probability that someone does not join a union conditional on not having been in a union in the prior time period (that is, continues their non-union status).
 - d. Using R or Python, reset your “prior_union” variable so that it takes on value 1 for all observations. For each observation, generate the predicted probability (or predicted value of the logit model) that someone is in a union and call it “pred11”. This prediction is the probability someone remains in a union (that is, continues their union status). Generate a variable called pred10 that is equal to $1 - \text{pred11}$. This prediction is the probability someone leaves a union conditional on having been in a union in the prior time period.
 - e. In a two-by-two table, find the average values for the four predictions created above in a manner consistent with that presented in class. You have now generated a model-based Markov matrix of transition probabilities (in this situation with respect to union status). Using R or Python, create such a table.
 - f. Following 1. above, use the Markov matrix calculated in 2e. to iterate the system forward until the Markov matrix has converged to its limit. What can you say about your long-term forecast of union membership among this sample of women? Are there actual statistics to which you could compare your results?
 - g. Submit all code and results.
3. In this exercise, you will work with Gaussian white noise random variables. [3points]
 - a. Generate and plot three Gaussian white noise random variables with 1,000 draws, the first with variance 1, the second with variance 2, the third with variance 4. (See HW 2.)
 - b. Generate two white noise random variables with 1,000 draws from a $N(0,1)$ distribution. Call them X1 and X2. Fit the bivariate linear model that relates X1 to X2 and report your regression results.
 - c. Recall your Monte Carlo simulation assignment in HW 2, Question 5. Using a similar approach, repeat b. above 1,000 times, each time recording the estimated value of the slope coefficient of the bivariate regression. Generate a histogram of your 1,000 replications.¹
 - d. Submit code and results.

¹ Not surprisingly, you will find that your histogram is somewhat bell-shaped and centered about zero. Increasing the number of replications will smooth the shape of the histogram.