

Assignment 1

Jacob Jameson

Due Feb 1

Collaborators: Jenna Rogers

Markov Models

Consider a Markov model with a cycle length of one year. For people who are well, their annual probability of becoming ill is 0.20 and their annual probability of dying from other causes is 0.05. For any given year (i.e., within a cycle), if people who are well become ill, they face a 0.30 probability that they will die from the illness that year; in subsequent years, their annual probability of dying from illness or other causes is 0.10. People who are ill can never return to the WELL state. Assume that people are initially well.

- (a) Write down a transition probability matrix for this Markov chain.

A transition probability matrix for this Markov chain can be defined in the following way:

$$\begin{bmatrix} 0.75 & 0.14 & 0.11 \\ 0 & 0.9 & 0.1 \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{array}{cc} & \begin{array}{ccc} \text{Well} & \text{Ill} & \text{Death} \end{array} \\ \begin{array}{c} \text{Well} \\ \text{Ill} \\ \text{Death} \end{array} & \begin{bmatrix} 0.75 & 0.14 & 0.11 \\ 0 & 0.9 & 0.1 \\ 0 & 0 & 1 \end{bmatrix} \end{array}$$

- (b) How many cycles must elapse by the time $\geq 25\%$ of the population is dead? Show all hand calculations using matrix algebra (you may use TreeAge software or Excel to check your calculations).

Assuming that all people are initially well, the current state of the population in $t = 0$ is:

$$\begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$$

To determine how many cycles must elapse before more than 25% of the population is dead, we just multiply our current state vector by the probability matrix in (a) as many times as we need until 25% of the population is dead.

$$\begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0.75 & 0.14 & 0.11 \\ 0 & 0.9 & 0.1 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 \times 0.75 & 1 \times 0.14 & 1 \times 0.11 \end{bmatrix} = \begin{bmatrix} 0.75 & 0.14 & 0.11 \end{bmatrix} \quad (1)$$

$$\begin{bmatrix} 0.75 & 0.14 & 0.11 \end{bmatrix} \begin{bmatrix} 0.75 & 0.14 & 0.11 \\ 0 & 0.9 & 0.1 \\ 0 & 0 & 1 \end{bmatrix} \quad (2)$$

$$\Rightarrow \begin{bmatrix} 0.75 \times 0.75 & 0.75 \times 0.14 + 0.14 \times 0.9 & 0.75 \times 0.11 + 0.14 \times 0.1 + 0.11 \times 1 \end{bmatrix} = \begin{bmatrix} 0.5625 & 0.231 & 0.2065 \end{bmatrix}$$

$$\begin{bmatrix} 0.5625 & 0.231 & 0.2065 \end{bmatrix} \begin{bmatrix} 0.75 & 0.14 & 0.11 \\ 0 & 0.9 & 0.1 \\ 0 & 0 & 1 \end{bmatrix} \quad (3)$$

$$\begin{bmatrix} 0.5625 \times 0.75 & 0.5625 \times 0.14 + 0.231 \times 0.9 & 0.5625 \times 0.11 + 0.231 \times 0.1 + 0.2065 \times 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0.421875 & 0.28665 & 0.291475 \end{bmatrix}$$

It takes 3 cycles before $\geq 25\%$ is dead.

```
# Define the starting population and probability matrix
pop <- matrix(c(1, 0, 0), nrow = 1, ncol = 3)

probability_matrix <- matrix(c(0.75, 0.14, 0.11,
                              0, 0.9, 0.1,
                              0, 0, 1), nrow = 3, byrow = TRUE)

# Simulate cycles until 25% or more dead
t = 0
while (pop[1,3] < 0.25){
  pop <- pop %*% probability_matrix
  t <- t + 1
}

print(paste('The number of cycles is ', t))
```

```
## [1] "The number of cycles is 3"
```

- (c) For persons who are ill, what is the 5-year survival probability? [Hint: consider your starting population]

To solve for the 5-year survival probability of people who are ILL, we can do this by initializing our starting population vector such that all of the people in the population are in the ILL category. We can then multiply our population vector by the transition probability matrix 5 times to simulate 5 cycles. Finally, the survival rate will be proportion of the population that is in the ILL state after the simulation. This is approximately 59.049%.

$$\begin{bmatrix} 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0.75 & 0.14 & 0.11 \\ 0 & 0.9 & 0.1 \\ 0 & 0 & 1 \end{bmatrix}^5 = \begin{bmatrix} 0 & 0.59049 & 0.40951 \end{bmatrix}$$

```
# Define the starting population and probability matrix
pop <- matrix(c(0, 1, 0), nrow = 1, ncol = 3)

probability_matrix <- matrix(c(0.75, 0.14, 0.11,
```

```

                                0, 0.9, 0.1,
                                0, 0, 1), nrow = 3, byrow = TRUE)

# Simulate 5 cycles
for (i in 1:5){
  pop <- pop %*% probability_matrix
}

print(paste('For persons who are ill, the 5-year survival probability is ',
            pop[1,2]*100, '%'))

```

```
## [1] "For persons who are ill, the 5-year survival probability is 59.049 %"
```

Consider the Markov model as described for parts (a)-(c) except that each year, people who are well face an unknown probability of becoming ill. A cohort has been followed for 5 years and the proportions of the cohort in the WELL, ILL, and DEAD states have been traced:

Cycle	WELL	ILL	DEAD
0	1.0000	0.0000	0.0000
1	0.6500	0.2100	0.1400
2	0.4550	0.3028	0.2422
3	0.3412	0.3362	0.3226
4	0.2730	0.3384	0.3886
5	0.2320	0.3237	0.4443

Table 1: Markov States

- (d) Use the Markov tracing to determine the annual probability of becoming ill. Does it remain constant over time?

Assuming the only probability that is unknown is the probability from WELL to ILL (and that all other transition probability are the same as (a)-(c)), we can deduce by looking at cycle 0 to cycle 1, that the probability of transitioning from WELL to ILL is 21%. That means that the transition state matrix is:

$$\begin{bmatrix} 0.65 & 0.21 & 0.14 \\ 0 & 0.9 & 0.1 \\ 0 & 0 & 1 \end{bmatrix}
 \quad
 \begin{array}{cc} & \begin{array}{ccc} \text{Well} & \text{Ill} & \text{Death} \end{array} \\ \begin{array}{c} \text{Well} \\ \text{Ill} \\ \text{Death} \end{array} & \begin{bmatrix} 0.65 & 0.21 & 0.14 \\ 0 & 0.9 & 0.1 \\ 0 & 0 & 1 \end{bmatrix} \end{array}$$

Now to see if that probability remains constant over time, we can simulate the 5 cycles and see if the ending state matrix is the same as that in the table.

```

# Define the starting population and probability matrix
pop <- matrix(c(1, 0, 0), nrow = 1, ncol = 3)

probability_matrix <- matrix(c(0.65, 0.21, 0.14,
                                0, 0.9, 0.1,
                                0, 0, 1), nrow = 3, byrow = TRUE)

table <- data.frame()
table <- rbind(table, pop)

```

```
# Simulate 5 cycles
for (i in 1:5){
  pop <- pop %*% probability_matrix
  table <- rbind(table, pop)
}
colnames(table) <- c('WELL', 'ILL', 'DEAD')
rownames(table) <- c('0', '1', '2', '3', '4', '5')

kable(table, caption = "Markov States")
```

Table 2: Markov States

	WELL	ILL	DEAD
0	1.0000000	0.0000000	0.0000000
1	0.6500000	0.2100000	0.1400000
2	0.4225000	0.3255000	0.2520000
3	0.2746250	0.3816750	0.3437000
4	0.1785063	0.4011788	0.4203150
5	0.1160291	0.3985472	0.4854238

As we can see, we get values that are slightly different than that in the original table, therefore we can conclude that the probability of going from WELL to ILL does not remain constant over time.

Consider a different Markov model with the following three states: WELL, ILL, and DEAD. Suppose that a new medication has been discovered, which can reduce the probability that illness develops by 50%. Tracing A (below) reflects the health of a population that is receiving prophylaxis with this new medication. A new treatment has also been discovered, which can reduce the mortality among those who become ill by 50%. Tracing B represents the health of the population that receives the new treatment once they become ill. Because of a deadly interaction between the medication and the treatment, only one option can be pursued. Assume that the utility of WELL is 1.0, the utility of ILL is 1.0, and the utility of DEAD is 0.

(e) Which strategy is preferred: new prophylaxis (Tracing A) or new treatment (Tracing B)?

The preferred strategy is the new treatment (Tracing B). We can conclude this from taking the sum of the reward column over the span of the total simulation. When doing this we see that the new treatment (Tracing B) has a higher total utility gained.

```
cycle.1 <-c(0.0000, 0.8726, 0.7383, 0.6071, 0.4857, 0.3735, 0.2744,
            0.1962, 0.1379, 0.0958, 0.0661, 0.0454, 0.0310, 0.0212,
            0.0144, 0.0098, 0.0067, 0.0045, 0.0000)

sum(cycle.1)
```

```
## [1] 3.9806
```

```
cycle.2 <- c(0.0000, 0.8762, 0.7430, 0.6103, 0.4863, 0.3747,
            0.2800, 0.2056, 0.1494, 0.1078, 0.0776, 0.0556,
            0.0398, 0.0285, 0.0204, 0.0146, 0.0104, 0.0074,
            0.0053, 0.0038, 0.0000)

sum(cycle.2)
```

[1] 4.0967

(f) What factor(s) could change the result noted in (e)?

Factors that could change the result noted in (e) could be if there was a smaller utility associated with the ILL state. Since new prophylaxis (Tracing A) allows people to transition from ILL to WELL, if the reward gain from WELL was higher that might allow new prophylaxis (Tracing A) to be the preferred strategy. Also since the new treatment (Tracing B) has a greater fraction of its lifetime cycle reward deriving from ILL people's utility, if that utility weight were smaller it would disproportionately effect the sum of the lifetime reward for Tracing B compared to Tracing A.

- (g) (Optional) A constant discount rate is used in the analysis. Determine this discount rate r from the tracings.

To determine the discount rate I will solve the equation:

$$CycleReward_n = \frac{state_n \cdot \vec{u}}{(1+r)^n}$$

Where $CycleReward_n$ is the cycle reward for the n th cycle, $state_n$ is the distribution of WELL, ILL, DEAD in n th cycle, \vec{u} is the utility weights defined in the problem, and r is the constant discount rate.

Solving this in cycle 1 for Tracing A:

$$0.8726 = \frac{0.9163}{(1+r)^1} \implies r = \frac{0.9163 - 0.8726}{0.8726}$$

$$\implies r \approx 0.05008022$$

We can validate this is the correct r by solving for every cycle reward of Tracing A and Tracing B:

```
states <- matrix(c(0.8788, 0.0375, 0.0838,
                  0.7513, 0.0627, 0.1860,
                  0.6245, 0.0783, 0.2972,
                  0.5043, 0.0860, 0.4097,
                  0.3833, 0.0934, 0.5233,
                  0.2731, 0.0949, 0.6323,
                  0.1946, 0.0814, 0.7240,
                  0.1386, 0.0650, 0.7963,
                  0.0988, 0.0499, 0.8514,
                  0.0704, 0.0373, 0.8923,
                  0.0501, 0.0274, 0.9224,
                  0.0357, 0.0200, 0.9443,
                  0.0255, 0.0145, 0.9601,
                  0.0181, 0.0104, 0.9715,
                  0.0129, 0.0075, 0.9796,
                  0.0092, 0.0054, 0.9854,
                  0.0066, 0.0038, 0.9896,
                  0.0047, 0.0027, 0.9926), nrow = 18, byrow = TRUE)

utility.weights <- c(1, 1, 0)
r <- 0.05
rewards <- rep(0,18)

for (cycle in 0:17){
  rewards[cycle+1] = round(sum(states[cycle+1,]*utility.weights)/(1+r)^cycle,4)
}

rewards

## [1] 0.9163 0.7752 0.6375 0.5099 0.3922 0.2883 0.2060 0.1447 0.1006 0.0694
## [11] 0.0476 0.0326 0.0223 0.0151 0.0103 0.0070 0.0048 0.0032
```