S&DS 238 Problem Set #7

Shaun Radgowski

October 30, 2020

Problem 1

(1a)

In order for $X_0, X_1, \dots$ to be a Markov chain, it would need to be true that X_{t+1} depends only on X_t. By manipulating the Y variables, we can prove that:

$$X_t = \frac{1}{2}(Y_{t-1} + Y_t),\ Y_t = 2X_t - Y_{t-1}$$

$$X_{t+1} = \frac{1}{2}(Y_t + Y_{t+1}) = \frac{1}{2}(2X_t - Y_{t-1} + Y_{t+1})$$

This violates the Markov property, as X_{t+1} is a function of additional random variables for different time steps (including $t-1$) beyond X_t. Thus, as the future is *not* independent of the past in this case, X_0, X_1, $\dots$ is not a Markov chain.

(1b)

$$X_t = \begin{cases} -1 & Y_{t-1} = -1, Y_t = -1 \\ 0 & Y_{t-1} = -1, Y_t = 1 \\ 0 & Y_{t-1} = 1, Y_t = -1 \\ 1 & Y_{t-1} = 1, Y_t = 1 \end{cases}$$

$$P\{X_{t+1} = -1 | X_t = -1\} = P\{Y_t = -1 | X_t = -1\} P\{Y_{t+1} = -1\} = \frac{1}{2}$$

$$P\{X_{t+1} = 0 | X_t = -1\} = P\{Y_t = -1 | X_t = -1\} P\{Y_{t+1} = 1\} = \frac{1}{2}$$

$$P\{X_{t+1} = -1 | X_t = 0\} = P\{Y_t = -1 | X_t = 0\} P\{Y_{t+1} = -1\}$$

$$P\{X_{t+1} = 0 | X_t = 0\} = P\{Y_t = -1 | X_t = 0\} P\{Y_{t+1} = 1\} + P\{Y_t = 1 | X_t = 0\} P\{Y_{t+1} = -1\}$$

$$P\{X_{t+1} = 1 | X_t = 0\} = P\{Y_t = 1 | X_t = 0\} P\{Y_{t+1} = 1\}$$

$$P\{X_{t+1} = 0 | X_t = 1\} = P\{Y_t = 1 | X_t = 1\} P\{Y_{t+1} = -1\} = \frac{1}{2}$$

$$P\{X_{t+1} = 1 | X_t = 1\} = P\{Y_t = 1 | X_t = 1\} P\{Y_{t+1} = 1\} = \frac{1}{2}$$

From this, it is clear that even when we know $X_t = 0$, we don't know if Y_t is 1 or -1. Because of this, to know $P\{X_{t+1} = x_{t+1}|X_t = 0\}$, it will still improve our knowledge to know X_{t-1}. One concrete example of this would be:

$$P\{X_{t+1} = 1|X_t = 0, X_{t-1} = 1\} \neq P\{X_{t+1} = 1|X_t = 0, X_{t-1} = -1\}$$

This is recursively true, where there are cases where knowing two time periods before will always improve the estimation of the current time period. This is also confirmed by the fact that Y_t extends back past $t = 0$ into negative values of t. For this reason, there is no single r value that will make the following statement true:

$$P\{X_{t+1} = x_{t+1}|X_t = x_t, ..., X_0 = x_0\} = P\{X_{t+1} = x_{t+1}|X_t = x_t, ..., X_{t-r+1} = x_{t-r+1}\}$$

Problem 2

(2a)

$$\pi_{t+1} = \pi_t P, \quad (\alpha \quad \beta) = (\alpha \quad \beta) \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix} = (\alpha(1-a) + b\beta \qquad a\alpha + \beta(1-b))$$

$$\alpha = \alpha(1-a) + b\beta, \quad \beta = a\alpha + \beta(1-b), \quad \beta(1-(1-b)) = a\alpha, \quad b\beta = a\alpha$$

We also know that $\alpha + \beta = 1$:

$$\alpha = 1 - \beta, \quad b\beta = a(1-\beta), \quad b\beta + a\beta = a$$

$$\beta = \tfrac{a}{b+a}, \quad \alpha = 1 - \beta = \tfrac{b}{b+a}$$

Thus, the stationary distribution $\pi = (\pi(1) \quad \pi(2)) = (\frac{b}{b+a} \quad \frac{a}{b+a})$.

(2b)

In this case, the Markov transition matrix is of the form $P = \begin{pmatrix} 0.85 & 0.15 \\ 0.1 & 0.9 \end{pmatrix}$, where the top left corner is the probability of going from True to True, the top right corner is the probability of going from True to False, the bottom left corner is the probability of going from False to True, and the bottom right corner is the probability of going from False to False.

The final answer in a very long examination is equal to:

$$\lim_{t\to\infty} \pi_t = \pi = \pi P = \pi \begin{pmatrix} 0.85 & 0.15 \\ 0.1 & 0.9 \end{pmatrix}$$

Where π is the stationary distribution for this matrix. If we match up this matrix with P from part a, we know that $a = 0.15$, $b = 0.1$. Using the solution from above, the stationary distribution is thus $\pi = (\frac{b}{b+a} \quad \frac{a}{b+a}) = (0.4 \quad 0.6)$. This shows that for large t (meaning a long examination), there is a 40% chance the last question will be True and a 60% chance the last question will be False. Accordingly, we would expect 40% of test items to be True and 60% to be False.

(2c)

We can simulate a very large number of examinations using R:

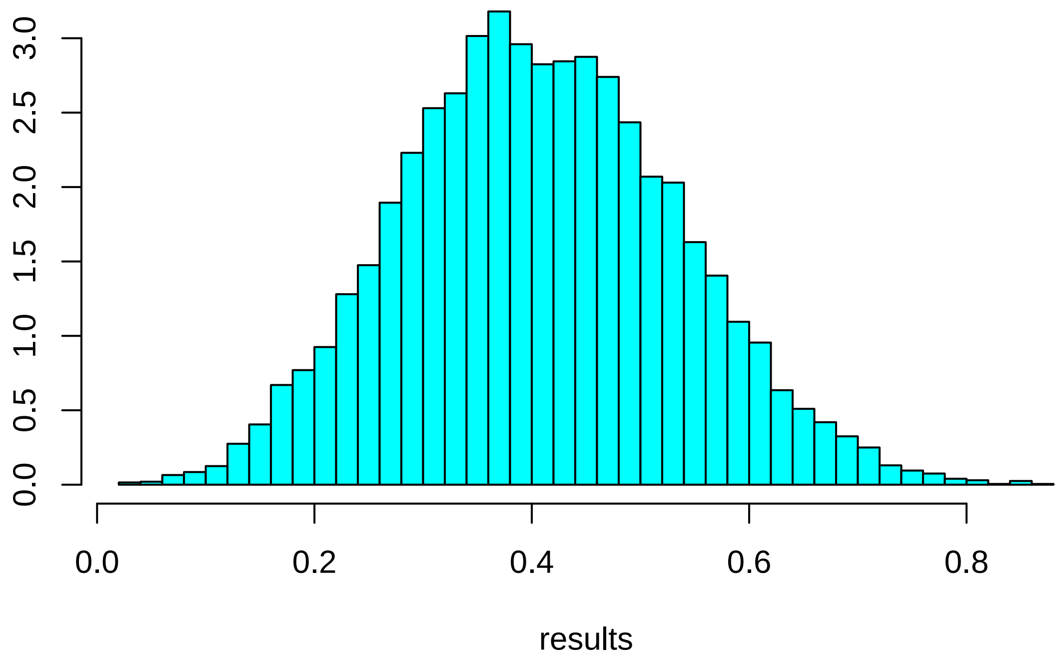
```r
next_question <- function(current_question) {
  random <- sample(1:100, 1)
  # If currently False:
  if (current_question == 0) {
    if (random < 11) {
      return(1)
    }
    else {
      return(0)
    }
  }
  # If currently True:
  else {
    if (random < 16) {
      return(0)
    }
    else {
      return(1)
    }
  }
}

take_test <- function(starting_value) {
  answers <- rep(NA, 100)
  answers[1] <- starting_value

  for (i in 1:99) {
    answers[i + 1] <- next_question(answers[i])
  }
  return(sum(answers)/100)
}

t <- 10000
results <- rep(NA, t)
for (i in 1:t) {
  starting <- sample(c(0, 1), 1)
  results[i] <- take_test(starting)
}

mean(results)
```

```
## [1] 0.402547
```

```r
sd(results)
```

```
## [1] 0.1273787
```

```r
library(MASS)
truehist(results)
```



(2d)

The precise expected value is $\frac{1}{100}\sum_{i=1}^{100} P\{I_i = 1\}$, where I_i indicates the event that question i is True. We can calculate the probability value for each I_i using matrix multiplication.

```r
matpow <- function(M, n) {
  ans <- diag(rep(1, nrow(M)))
  for (i in 1:n) {
    ans <- ans %*% M
  }
  return(ans)
}

P <- rbind(c(0.85, 0.15), c(0.1, 0.9))
total <- 0.5
for (i in 1:99) {
  probs <- c(0.5, 0.5) %*% matpow(P, i)
  total <- total + probs[1,1]
}
total / 100
```

```
## [1] 0.404
```

Problem 3

(3a)

$$P = \begin{pmatrix} P\{1 \to 1\} & P\{1 \to 2\} & P\{1 \to 3\} \\ P\{2 \to 1\} & P\{2 \to 2\} & P\{2 \to 3\} \\ P\{3 \to 1\} & P\{3 \to 2\} & P\{3 \to 3\} \end{pmatrix} = \begin{pmatrix} (1-p)\frac{1}{3} & (1-p)\frac{1}{3} + \frac{p}{2} & (1-p)\frac{1}{3} + \frac{p}{2} \\ (1-p)\frac{1}{3} & (1-p)\frac{1}{3} & (1-p)\frac{1}{3} + p \\ (1-p)\frac{1}{3} + p & (1-p)\frac{1}{3} & (1-p)\frac{1}{3} \end{pmatrix} = \begin{pmatrix} 0.1 & 0.45 & 0.45 \\ 0.1 & 0.1 & 0.8 \\ 0.8 & 0.1 & 0.1 \end{pmatrix}$$

(3b)

Solving the system of equations by hand:

$$\begin{pmatrix} a & b & c \end{pmatrix} = \begin{pmatrix} a & b & c \end{pmatrix} \begin{pmatrix} 0.1 & 0.45 & 0.45 \\ 0.1 & 0.1 & 0.8 \\ 0.8 & 0.1 & 0.1 \end{pmatrix} = \begin{pmatrix} 0.1(a+b) + 0.8c & 0.45a + 0.1(b+c) & 0.45a + 0.8b + 0.1c \end{pmatrix}$$

$$0.9a = 0.1b + 0.8c, \quad 0.9b = 0.45a + 0.1c, \quad 0.9c = 0.45a + 0.8b$$

And because we know that $a + b + c = 1, \quad a = 1 - b - c$:

$$a = \frac{0.1b + 0.8c}{0.9} = 0.11b + 0.88c = 1 - b - c, \quad b = \frac{1 - 1.88c}{1.11} = 0.9 - 1.7c$$

$$0.9(0.9 - 1.7c) = 0.45a + 0.1c, \quad c = \frac{0.81 - 0.45a}{1.63} = 0.5 - 0.28a$$

$$0.9(0.5 - 0.28a) = 0.45a + 0.8(0.9 - 1.7(0.5 - 0.28a)), \quad a = 0.38$$

$$c = 0.5 - 0.28(0.38) = 0.39, \quad b = 0.9 - 1.7(0.39) = 0.23$$

With this method, we have found $\pi = \begin{pmatrix} \pi(1) & \pi(2) & \pi(3) \end{pmatrix} \approx \begin{pmatrix} 0.38 & 0.23 & 0.39 \end{pmatrix}$

Instead, solving by raising the probability matrix to a high power:

```r
matpow <- function(M, n) {
  ans <- diag(rep(1, nrow(M)))
  for (i in 1:n) {
    ans <- ans %*% M
  }
  return(ans)
}

P <- rbind(c(0.1, 0.45, 0.45), c(0.1, 0.1, 0.8), c(0.8, 0.1, 0.1))
c(0.333, 0.333, 0.333) %*% matpow(P, 100)
```

```
##          [,1]      [,2]      [,3]
## [1,] 0.374946 0.2311311 0.3929229
```

We get the same answer.

(3c)

PageRank ranks Page 3 highest, Page 1 second-highest, and Page 2 last.

Problem 4

(4a)

If we enumerate the set options HH, HT, TH, TT as 1, 2, 3, and 4 (respectively), then given that the most recent pair of tosses is one of the set, the probability of the next pair of tosses being one of these is given by the probability transition matrix:

$$P = \begin{pmatrix} P\{1 \to 1\} & P\{1 \to 2\} & P\{1 \to 3\} & P\{1 \to 4\} \\ P\{2 \to 1\} & P\{2 \to 2\} & P\{2 \to 3\} & P\{2 \to 4\} \\ P\{3 \to 1\} & P\{3 \to 2\} & P\{3 \to 3\} & P\{3 \to 4\} \\ P\{4 \to 1\} & P\{4 \to 2\} & P\{4 \to 3\} & P\{4 \to 4\} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

This makes sense, because if the second flip of the current pair is Heads (like in S 1 or 3), then the only pairs possible for the next element in the chain are the pairs where the first flip is Heads (like S 1 or 2). The same is true in reverse, for Tails.

(4b)

Using the Law of Total Expectation:

$$E_i(\tau_j) = \textstyle\sum_k P_i\{X_1 = k\} E(\tau_j | X_1 = k)$$

Because we know that the $X_0 = i$, the first element in that sum (the probability of getting $X_1 = k$) is simply the probability transition matrix element $P(i,k)$. The second element in the sum (the expected value given that $X_1 = k$) is simply $1 + E_k(\tau_j)$. This is because if $E_k(\tau_j)$ is equivalent to starting over to $E_i(\tau_j)$ where this time $i = j$, but now 1 time unit has been added to the expected value. In either case, $1 + E_k(\tau_j)$ describes the timeless property that you have just added one time unit to the expected value but you are not necessarily one step closer.

From there, we have:

$$E_i(\tau_j) = \textstyle\sum_k P(i,k)(1 + E_k(\tau_j)) = \sum_k P(i,k) + P(i,k)E_k(\tau_j))$$

$$= \textstyle\sum_k P(i,k) + \sum_k P(i,k)E_k(\tau_j)) = 1 + \sum_k P(i,k)E_k(\tau_j))$$

Where we have used the Law of Total Probability to state that $\sum_k P(i,k) = 1$. This can also be confirmed by summing accross rows in the matrix above, where each row sums to 1.

(4c)

$$E_i(N_j) = 1 + \textstyle\sum_k P(i,k)E_k(N_j), \;\; E_1(N_1) = 0$$

$$E_2(N_1) = 1 + \textstyle\sum_k P(2,k)E_k(N_1) = 1 + \frac{1}{2}E_3(N_1) + \frac{1}{2}E_4(N_1)$$

$$E_3(N_1) = 1 + \textstyle\sum_k P(3,k)E_k(N_1) = 1 + 0 + \frac{1}{2}E_2(N_1)$$

$$E_4(N_1) = 1 + \textstyle\sum_k P(4,k)E_k(N_1) = 1 + \frac{1}{2}E_3(N_1) + \frac{1}{2}E_4(N_1)$$

$$E_2(N_1) = 1 + \tfrac{1}{2}(1 + \tfrac{1}{2}E_2(N_1)) + \tfrac{1}{2}E_2(N_1) = \tfrac{3}{2} + \tfrac{3}{4}E_2(N_1)$$

$$E_2(N_1) = 6, \quad E_4(N_1) = E_2(N_1) = 6, \quad E_3(N_1) = 1 + \frac{1}{2}E_2(N_1) = 4$$

Thus, using the Law of Total Expectation, we know that $E(N_j) = 2 + \sum_k P\{X_0 = k\}E_k(N_j)$, where the extra 2 is added for the first two flips that constitute X_0. Using this and the fact that all four starting patterns have equal probability of $\frac{1}{4}$, we can show that $E(N_1) = 2 + \frac{1}{4}0 + \frac{1}{4}6 + \frac{1}{4}4 + \frac{1}{4}6 = 6$.

(4d)

Using the same logic as part c:

$$E_i(N_j) = 1 + \sum_k P(i,k)E_k(N_j), \;\; E_2(N_2) = 0$$

$$E_1(N_2) = 1 + \sum_k P(1,k)E_k(N_2) = 1 + \frac{1}{2}E_1(N_2) + 0$$

$$E_3(N_2) = 1 + \sum_k P(3,k)E_k(N_2) = 1 + \frac{1}{2}E_1(N_2) + 0$$

$$E_4(N_2) = 1 + \sum_k P(4,k)E_k(N_2) = 1 + \frac{1}{2}E_3(N_2) + \frac{1}{2}E_4(N_2)$$

$$E_1(N_2) = 2(1) = 2, \quad E_3(N_2) = E_1(N_2) = 2$$

$$E_4(N_2) = 2(1 + \frac{1}{2}(2)) = 4$$

Thus, using the Law of Total Expectation, we again know that $E(N_j) = 2 + \sum_k P\{X_0 = k\}E_k(N_j)$, where the extra 2 is added for the first two flips that constitute X_0. Using this and the fact that all four starting patterns have equal probability of $\frac{1}{4}$, we can show that $E(N_2) = 2 + \frac{1}{4}2 + \frac{1}{4}0 + \frac{1}{4}2 + \frac{1}{4}4 = 4$.

Problem 5

```
expected_value <- function(j) {
  # j is the combination of interest (vector)
  results = sample(c(0, 1), 2, replace=TRUE)
  tosses = 2
  while (TRUE) {
    if (identical(results, j)) {
      return(tosses)
    }

    results[1] = results[2]
    results[2] = sample(c(0, 1), 1)
    tosses = tosses + 1
  }
}

simulation <- function(j) {
  n <- 10000
  # n = number of simulations to run
  overall_results = rep(NA, n)
  for (i in 1:n) {
```

```r
    overall_results[i] = expected_value(j)
  }
  return(mean(overall_results))
}

# 1 is Heads, 0 is Tails
# N1: Heads, Heads
simulation(c(1, 1))
```

```
## [1] 6.0388
```

```r
# N2: Heads, Tails
simulation(c(1, 0))
```

```
## [1] 4.0129
```

This simulation confirms the results from Problem 4.