# Stochastic Processes
## ASSIGNMENT ONE

**Santiago Alfonso Raposo Briceño - 100414456**
**Javier Martínez Llamas - 100392684**
**Ricardo Hortelano Sánchez - 100418220**

## 1 Problem 1

### 1.1 Question a

We can model the different channels as states in a Markov chain, with two absorbing states representing the conversion or non-conversion of a client. This assumes however that the channel each user explores is only decided by the last channel he or she clicked on, which might not be completely realistic.

### 1.2 Question b

The state space is all of the unique values of the chain existing in our sample. This is easily computed with R, and it appears our state space consists of:

- **Three channels** labeled as `Ch 1`, `Ch 2` and `Ch 3`.

- **Two absorbing states** representing the `Conversion` and `Non-conversion` states.

We found the estimate of the transition probabilities using maximum likelihood, as we have from our notes:

$$\hat{p}_{ij} = \frac{n_{ij}}{\sum_{j=1}^{k} n_{ij}} \quad i, j \in S$$

Where $n_{ij}$ is the total number of transitions from state $i$ to state $j$ and the denominator is the total amount of transitions from state $i$ to any other state $j$. The estimate found for our sample is the following transition matrix:

$$P = \begin{bmatrix} 0.3042 & 0.0792 & 0.1518 & 0.3132 & 0.1515 \\ 0.1660 & 0.1612 & 0.3365 & 0.1601 & 0.1761 \\ 0.1358 & 0.0756 & 0.1459 & 0.2202 & 0.4225 \\ 0.0000 & 0.0000 & 0.0000 & 1.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 1.0000 \end{bmatrix},$$

where we have encoded our states as numbers using the following mapping: `Ch 1` is row (and column) 1, `Ch 2` is number 2, `Ch 3` is number 3, `Conversion` is number 4 and `Non-Conversion` is number 5.

### 1.3 Question c

Our first consumer journey is `Ch 3 -> Ch 2 -> Conversion`, which is equivalent in our numerical mapping to $X_1 = 3, X2 = 2, X3 = 4$. The probability we then have to calculate is $P(X_3 = 4|X_2 = 2, X_1 = 3)$. This can be done as follows:

$$P(X_3 = 4|X_2 = 2, X_1 = 3) = P(X_3 = 4|\ X_2 = 2) * P(X_2 = 2|X_1 = 3) * P(X_1 = 3)$$
$$= P_{2,4} * P_{3,2} * (\boldsymbol{\alpha}P)_3$$

Assuming our initial distribution $\boldsymbol{\alpha} = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0, 0)$[1] the answer is 0.0026.

---

[1] We are assuming the client cannot start in one of the conversion or non-conversion states and that it is uniform on the rest, which seems reasonable.

## 1.4  Question d

We can find the limiting distribution for the chain by using the result from equation 3.11 in Dobrow. This allows us to compute the limiting submatrix as $(\boldsymbol{I} - \boldsymbol{Q})^{-1}\boldsymbol{R}$ for the absorbing states, which is:

$$(\boldsymbol{I} - \boldsymbol{Q})^{-1}\boldsymbol{R} = \begin{bmatrix} 0.5887 & 0.4113 \\ 0.4650 & 0.5350 \\ 0.3925 & 0.6075 \end{bmatrix}$$

Where the first column represents the limiting distribution of the conversion state and the second column represents the limiting distribution of the non-conversion state. We can then compute the total conversion ratio as the ratio between the sum of column 1 (which is the total proportion of realizations that end in a conversion) divided by the sum of the matrix.

$$\text{Conversion Ratio } = 0.4821.$$

## 1.5  Question e

We get the following results for the two approaches to calculating the removal effect.

| Approach | State | Removal Effect |
|---|---|---|
| Completely removing the state | Ch 1 | 0.2085 |
| | Ch 2 | -0.0285 |
| | Ch 3 | -0.2328 |
| Making the state "useless" | Ch 1 | 0.5921 |
| | Ch 2 | 0.4013 |
| | Ch 3 | 0.4681 |

Table 1: Removal effect of both approaches

The conversion ratio is defined as $1 - \frac{CR_{new}}{CR_{old}}$. We can see then that if the new conversion ratio is smaller than the old conversion ratio, the removal effect will be positive and it will be negative otherwise. Seeing table 1 we can conclude that:

- **Channel 1** always has a positive removal effect, which means that removing it causes the new conversion rate to be lower than the global conversion rate. Both approaches then seem to indicate that removing channel one would be counterproductive, as it is contributing to a high amount of conversions.

- **Channel 2** does not seem to change the conversion rate much with the first approach, but for the second approach it does make the new conversion rate lower. Therefore we can assume that this state is maybe not as important as the first one, but it does have a positive impact on the conversion rate.

- The removal of **Channel 3** actually has contradicting results for the two strategies. In the first one, removing it actually raises the conversion rate, but with the second approach its removal decreases the conversion rate by almost half. We would have to further examine why this is the case to conclude whether the removal of this channel would affect the conversion rate positively or not.

The main difference between the two approaches is that in the first one, as we are removing a state (row and column) from the transition matrix, we have to normalize it so that the rows

sum to 1 again, while in the second approach we just change the probabilities for the state we are studying. So in some sense in the first approach we are artificially "inflating" the other transition probabilities.

# 2 Question 2

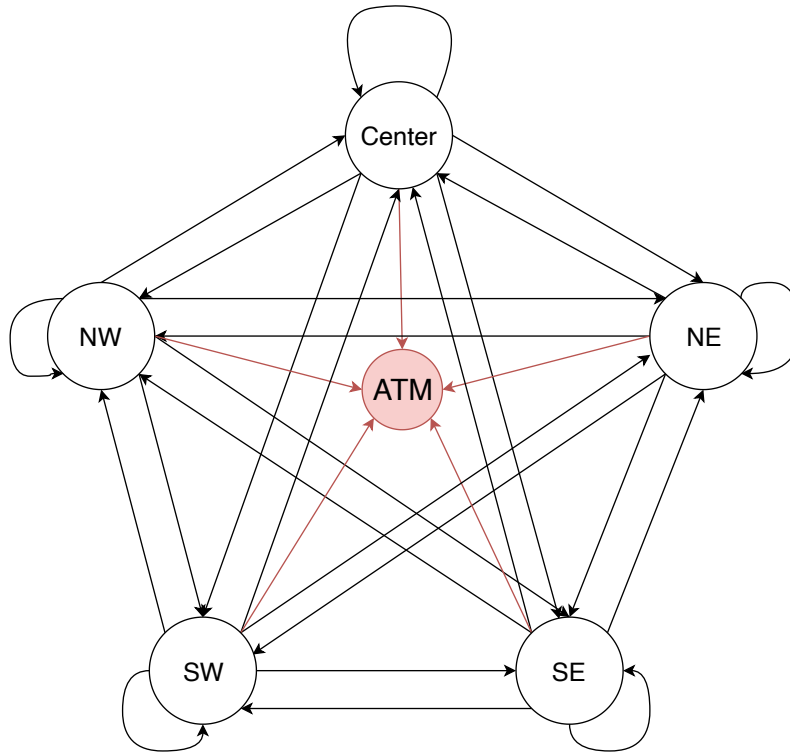## 2.1 Question a

We can see the transition graph in fig. 1.



Figure 1: Transition Graph

The state `Atm` represents a particle escaping into the atmosphere, and it is the only absorbing state. The other states are all transient and aperiodic (as they all have loops to themselves). Also, there is just one communication class in the chain, as all states communicate.

We have not represented the probabilities in the graph as they are not given and it would be quite noisy. The arrows indicate that in principle the states are all connected with each other (in both directions).

## 2.2 Question b

### Part I

Let $\sum_j q_{ij}$ be the probability of particle transition from $i$ to other areas and $q_i$ the particles escaping to the atmosphere, where $\sum_j q_{ij} + q_i = 1$. So that the total transition of particles to area $j$ is given by $\sum_i q_{ij}$.

Let the total number of particles present on area $j$ on day $n$ be a random variable $W_j^n$. And let the number of particles transitioning from area $i$ to area $j$ on day $n$ be a random variable

$B_{ij}^n$. Then, the expected total number of particles transited to area $j$ will be given by

$$\mathbb{E}[B_{ij}^n] = \sum_i q_{ij}\mathbb{E}[W_i^{n-1}].$$

Or the total particles on area $i$ weighted with the probability of transition to area $j$. Therefore, this can be reformulated as the success that a particle transits from area $i$ to area $j$, such that

$$B_{ij}^n \sim Bin(W_i^{n-1}, q_{ij}).$$

An area $j$ emits every day a number of particles given by the random variable $U_j$. So that the total number of particles on day $n$ on area $j$, $W_j^n$, will be given by the particles emitted that day plus the particles of the previous day transited from other areas. That is

$$W_j^n = U_j + \sum_i B_{ij}^n$$

We can then conclude that the expected number of particles on day $n$ present on area $j$ is

$$\mathbb{E}[W_j^n] = \mathbb{E}[U_j] + \sum_i q_{ij}\mathbb{E}[W_i^{n-1}].$$

**Part II**

Let $X$ be a random variable and let $A_1, \ldots, A_k$ be a sequence of events that partition the sample space. Then,

$$\mathbb{E}[X] = \sum \mathbb{E}[X|A_i]P(A_i)$$

In the same way, let $Y$ be any random variable on the same probability space than $X$. Then,

$$\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X|Y]]$$

Given the Markov Chain problem previously defined, where $w_i^n$ is the expectation of the random variable $W_i^n$ or the expected total number of particles present on area $i$ on day $n$. Since $W_i^n$ and $W_i^{n-1}$ follow the same distribution, then

$$\mathbb{E}[W_i^n] = \mathbb{E}[\mathbb{E}[W_i^n|W_i^{n-1}]],$$

and it follows that

$$w_j^n = u_j + \sum_i q_{ij}w_i^{n-1}$$
$$\mathbf{w}^n = \mathbf{u} + \mathbf{Q}\mathbf{w}^{n-1}.$$

**Part III**

Given
$$\mathbf{w}^n = \begin{bmatrix} w_1^n \\ w_2^n \\ w_3^n \\ w_4^n \\ w_5^n \end{bmatrix} \qquad \mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \end{bmatrix} \qquad \mathbf{Q} = \begin{bmatrix} q_{11} & \cdots & q_{15} \\ \vdots & \ddots & \vdots \\ q_{51} & \cdots & q_{55} \end{bmatrix},$$

show that the expected total number of particles on day $n$ is equal to the sum of emitted particles weighted with the transition probabilities.

$$\mathbf{w}^n = \mathbf{u} + \mathbf{uQ} + \cdots + \mathbf{uQ^{n-1}} = \mathbf{u}\sum_{i=0}^{n-1}\mathbf{Q}^i$$

Knowing that the expected number of particles is given by

$$\mathbb{E}[W_j^n] = \mathbb{E}[U_j] + \sum_i q_{ij}\mathbb{E}[W_i^{n-1}]$$

$$w_j^n = u_j + \sum_i q_{ij}w_i^{n-1}$$

$$\mathbf{w}^n = \mathbf{u} + \mathbf{Qw}^{n-1},$$

we can see that, for the initial state $n = 0$ and without any transitioning of particles , the statement is true such that $\mathbf{w}^0 = \mathbf{u}$ , where $\mathbf{w}^{-1} = 0$

Assuming that for any state $n \in \mathbb{N}$ the statement is true we prove that for state $n+1$ it is also true.

$$\mathbf{w}^{n+1} = \mathbf{u} + \mathbf{Qw}^n = \mathbf{u} + \mathbf{Q}(\mathbf{u} + \mathbf{Qw}^{n-1})$$

$$= \mathbf{u} + \mathbf{uQ} + \mathbf{Q}^{n+1-(n-1)}\mathbf{w}^{n-1} = \ldots$$

$$= \mathbf{u}\sum_{i=0}^{n}\mathbf{Q}^i$$

## 2.3 Question c

We are asked to show that $\mathbf{w}^n \xrightarrow[n\to\infty]{} \mathbf{w}$ and show how to compute $\mathbf{w}$ from $\mathbf{u}$.

Knowing that $\mathbf{w}^n = \mathbf{u}\sum_{i=0}^{n-1}\mathbf{Q}^i$ we want to study the long term behavior of $\mathbf{w}^n$ such that

$$\lim_{n\to\infty}\mathbf{w}^n = \lim_{n\to\infty}\mathbf{u}\sum_{i=0}^{n-1}\mathbf{Q}^i = \mathbf{u}\sum_{i=0}^{\infty}\mathbf{Q}^i$$

Since all states of $\mathbf{Q}$ are transient and $0 \leq \sum_j q_{ij} < 1 \; \forall i$ we know that $\mathbf{Q}^n$ converges to $\mathbf{0}$ as $n \to \infty$. Then,

$$\sum_{i=0}^{\infty}\mathbf{Q}^i = (\mathbb{I} - \mathbf{Q})^{-1}$$

Therefore,

$$\mathbf{w} = \lim_{n\to\infty}\mathbf{w}^n = \mathbf{u}\sum_{i=0}^{\infty}\mathbf{Q}^i = \mathbf{u}(\mathbb{I} - \mathbf{Q})^{-1}.$$

## 2.4 Question d

In order to limit the pollution levels and the amount of $PM_{10}$ ($W_j$), regardless of the possible transitions of particles, constant over time and beyond the scope of the city council, the only way to achieve this limit is, as previously seen, by the random variable $U_i$ that defines the emissions on each area $i$ such that

$$W_j^n = U_j + \sum_i B_{ij}^n$$

Therefore, depending on the distribution of $U_j$ and, consequently, its expectation, the number of particles could be controlled and prescribed.