The Scientific Journal of Business and Finance

https://caf.journals.ekb.eg

Statistical Analysis of Internal Migration in Egypt Using Markov Chains

Asmaa Antar [a] Mervat M. Elgohary [a] and Amira Abd Elghany [a]

[a] Department of Statistics, Faculty of Commerce, Al-Azhar University, Cairo, Egypt

Published online: **March 2024.**

To cite this article: Antar, Asmaa & El-Gohary, Mervat and Abd Elghany, Amira. Statistical Analysis of Internal Migration in Egypt Using Markov Chains , The Scientific Journal of Business and Finance, 44 (1), 360-379

DOI: https://doi.org/10.21608/caf.2024.351771

*Corresponding author: zeinab.nawar@eui.edu.eg

Statistical Analysis of Internal Migration in Egypt Using Markov Chains

Asmaa Antar [a] Mervat M. Elgohary [a] and Amira Abd Elghany [a]

Department of Statistics, Faculty of Commerce, Al-Azhar University, Cairo, Egypt

Article History

Received 4 February 2024, Accepted 18 February 2024, Available online March 2024

Abstract

This paper aims to analyze internal migration in Egypt using Markov Chain method to predict the numbers of migrants for each region of Egypt for the coming periods. This study aims to find the matrix of transition probabilities that shows the probabilities of migrants moving between the regions, and to find the mean first passage time and the mean recurrence time. The states of the Markov chain model are Cairo region, Alexandria region, Delta region, Canal region, North Upper Egypt region, Central Upper Egypt region, and Southern Upper Egypt region these states are communicated to each other and irreducible. The results of the analysis; show that the Cairo region is the only region that attracts migrants, whether male or female, while the rest of the regions are regions that expel people, and the rate of expulsion varies in each of them. Moreover, it is found that the North Upper Egypt region is the most region that expells the migrants, especially to Cairo region, and it is also the least region that retains migrants, whether male or female. We conclude from the forecasting process that the percentage of migrants, whether male or female, for the Cairo region will increase, and the percentage of migrants for the rest of the regions of the Republic will decrease.

Keywords: Markov chain, steady state, internal migration, transition probability matrix, the equilibrium matrix, initial probability, stationary probability

1. Introduction

The phenomenon of internal migration is one of the most important phenomena in the field of geographic studies in general and population geography studies in particular, as migration represents one of the three components responsible for population change in any region, namely deaths, births, and migration, but migration is the most influential of these components, Then the phenomenon of internal migration in Egypt is considered one of the demographic phenomena that must be studied with care and attention for two reasons. The first reason is its geographical nature, dividing the governorates into urban and rural areas, and the obvious differences in the processes of development between the governorates of Egypt. The second reason is that it is based on several economic and social factors on the one hand, and leads to many changes in the age and gender composition of the population on the other hand. This was the motive behind the study of internal migration in Egypt. Internal migration is defined as the movement of an individual or a group of individuals from one geographical unit to another within the borders of one state, whether that includes a permanent or temporary change in the place of residence **(حسانين, 2009).**

The migration of people from one region to another is considered a random phenomenon that can be studied and analyzed through stochastic models to determine the changes that occur in the population structure and labor force in the origin (departure) and arrival (destination) regions and the reasons behind the migration process. There are many studies that have provided an analysis of the phenomenon of internal migration using probabilistic models, such as the Markov chains models, including a study of (Compton, 1968, Hierro, 2008, Gulshan, 2015, Usman et al., 2015, Garrocho et al., 2016 Venkatesan and Sasikala, 2018, Yip et al., 2020).

The aim of this paper is to study the internal migration in Egypt by using Markov Chain method, predict the numbers of migrants for each region of Egypt for the coming periods, which helps the state in setting and developing development plans among the regions of Egypt. The data used in this paper are from the Central Agency for Public Mobilization and Statistics (CAPMAS), according to the latest census of 2017, which include the number of migrants between the governorates of Egypt according to the governorate of previous and current residence. This paper is organized as follows: Section 2 presents Markov chains. Section 3 illustrates the results of analysis of internal migration data in Egypt according to the 2017 Census. Finally, section 4 presents conclusions.

2. Markov chain

Markov chain is a method of analyzing the current movement of a phenomenon in an attempt to predict the future movement of the same phenomenon. This method dates back to the Russian scientist *Andrei Markov* in the early twentieth century (Agbinya, 2020).

Markov chain is a special type of stochastic processes{X(t), t ∈ T} where X(t) is a random variable, the values that a random variable takes are called states, therefore the set that includes all possible values of states is called the state space and is denoted by symbol (S).The state space may be discrete or continuous. If the state space is discrete, the process is referred to as a chain and the states are usually identified with the set of natural numbers S = {0,1,2,...} or a subset of it (Stewart,2009).
T is known as a parameter space or index set and is a subset of $\{-\infty, +\infty\}$, this set includes all possible values for the parameter (t). If the parameter space is discrete, e.g., T = {0,1,2,…} then the stochastic process {X(n), n ∈ T} is called a discrete- time parameter stochastic process. If the parameter space is continuous, e.g., T = {t: $0 \leq t < +\infty$}, then the stochastic process {X(t), t ∈ T} is called a continuous- time parameter stochastic process.
A stochastic process {X(n), n ∈ T} with a Markov property and discrete- time discrete- state is called a discrete- time discrete- state Markov chain. The basic property of a Markov chain is simply that for the process, the future and past states of the process are independent if its present state is known this is called the Markov property. It means that X_{n+1} depends upon X_n, but it does not depend upon X_{n-1},…, X_1, X_0. The mathematical formula for the Markov property is as follows:

$$P_r(X_{n+1} = j|\, X_n = i, X_{n-1} = i_{n-1}, \dots, X_1 = i_1, X_0 = i_0) = P_r(X_{n+1} = j|X_n = i) \quad (1)$$

for all states $i_0, i_1, \dots, i_{n-1}, i, j$ and for all n.

(Agbinya, 2020, Oliver, 2009)

The conditional probabilities P_r $(X_{n+1} = j|X_n = i)$ are known as "one- step transition probabilities" p_{ij} , which means the probability of transition from state i to state j, where;

$$P_{ij} = P_r(X_{n+1} = j|X_n = i).$$

Transition probabilities are listed in square matrix, the matrix is called the "transition probability matrix" and is denoted with the symbol (P) as follows:

$$P = P_{ij} = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1m} \\ p_{21} & p_{22} & \cdots & p_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ p_{m1} & p_{m2} & \cdots & p_{mm} \end{bmatrix} \qquad (2)$$

Where,

1) $p_{ij} \geq 0$

2) $\sum_{j=1}^{m} p_{ij}$=1, i, j = 1,2,…,m.

This matrix is called "one- step transition probability matrix". If we want to obtain the matrix of transitional probabilitiesP^n after n steps or n time periods it can be written as follows:

$$P^n = P_{ij}^n = \begin{bmatrix} P_{11}^n & P_{12}^n & \cdots & P_{1m}^n \\ P_{21}^n & P_{22}^n & & P_{2m}^n \\ \vdots & \vdots & \ddots & \vdots \\ P_{m1}^n & P_{m2}^n & \cdots & P_{mm}^n \end{bmatrix} \qquad (3)$$

We can get this matrix by multiplying the original matrix (P) by itself (n) times, as follows:

$$P^2 = P * P, P^3 = P^2 * P, \ldots$$

2.1. **Initial probability distribution**

The vector $\pi^{(0)} = \{\pi_1^{(0)}, \pi_2^{(0)}, \ldots \pi_m^{(0)}\}$ is called the initial probability vector or initial probability distribution of a Markov chain $\{X_n, n = 0,1,2, \ldots\}$ with state space S ={1,2,3,…,m}, where $\pi_j^{(0)} = P_r(X_0 = j), j \in S$ if the following conditions are met:

1) $\pi_j^{(0)} \geq 0 \quad \forall\, j \in S$,

2) $\sum_{j \in S} \pi_j^{(0)}$=1.

This probability, $\pi_j^{(0)} = P_r(X_0 = j)$, $j \in S$ means that the process will start from state $j \in S$ with probability $\pi_j^{(0)}$. Hence, it can be understood that the probability vector $\pi_j^{(0)}$ gives the probability distribution of the Markov chain at the beginning before the first transition.

The probability of a stochastic process in state $j \in S$ after (n) transitions is denoted by the symbol $\pi_j^{(n)}$ and is defined by the following relationship:

$$\pi_j^{(n)} = P_r(X_n = j), j \in S. \qquad (4)$$

Given the initial probability vector and transition probability matrix, we can get the probability vector after (n) steps as follows:

$$\pi^{(n)} = \pi^{(0)} * P^{(n)}$$

Or

$$\pi^{(n)} = \pi^{(n-1)} * P$$

2.2. Stationary probability distribution

Definition: Let P be the transition matrix of a Markov chain with state space S. A probability distribution $\pi = (\pi_1, \pi_2, \dots)$ on S satisfying $\pi P = \pi$ is called a stationary distribution of the chain.

The entries of π thus satisfy

$$\pi_j = \sum_{i \epsilon S} p_{ij} \pi_i \qquad for\ all\ j\ \epsilon\ S \tag{5 }$$

$$\sum_{i \epsilon S} \pi_i = i$$

2.3. Steady-state of regular Markov chain

The steady state of the Markov chain appears when the stochastic process continuous for long periods of time. The long-run probability is also called the steady-state probability.

Definition: Let P_{ij}^n be the n-step transition probabilities of a regular Markov chain. If there exists a probability distribution π on S such that

$$P_{ij}^n \rightarrow \pi_j \ \ as\ n \rightarrow \infty \ \ for\ all\ i, j \ \in S$$

We call $\pi_j$is the limit distribution of the Markov chain when the n-step transition matrix converges to a limit matrix in which all rows are identical and its shape is as follows:

$$\lim_{n \to \infty} P^{(n)} = \hat{P}$$

$$= \begin{bmatrix} P_1 & P_2 & P_3 & \dots \\ P_1 & P_2 & P_3 & \dots \\ P_1 & P_2 & P_3 & \dots \end{bmatrix} \tag{6}$$

The $\hat{P}$ is a matrix whose rows are equal to the stationary distribution or limit distribution π_j.

Where

$$\lim_{n\to\infty} P^{(n)} = \hat{P} = \pi_j \tag{7}$$

The limit distribution π_j describes the probability that the chain is in state j at some late time-point and that at this time, the chain has "forgotten how it started" (Olofsson and Andersson, 2012).

2.4. Mean first passage and recurrence times

For Ergodic Markov chain (if all states in a Markov chain are recurrent, aperiodic, and communicate with each other, the chain is said to be ergodic), let μ_{ij} denotes the expected number of transitions or mean first passage time before reaching state j, provided that the chain is currently in state i, which is defined by:

$$\mu_{ij} = \begin{cases} \infty & if\ \sum_{n=1}^{\infty} f_{ij}^{(n)} < 1 \\ \sum_{n=1}^{\infty} n f_{ij}^{(n)} & if\ \sum_{n=1}^{\infty} f_{ij}^{(n)} = 1 \end{cases} \tag{8}$$

μ_{ij} uniquely satisfies the equation

$$\mu_{ij} = p_{ij}^{(1)} + \sum_{k\neq j} p_{ik}(1 + \mu_{kj}) \tag{9}$$

Where

$$p_{ij} + \sum_{k\neq j} p_{ik} = 1$$

Equation (9) can be rewritten in the following form:

$$\mu_{ij} = 1 + \sum_{k\neq j} p_{ik}\mu_{kj}. \tag{10}$$

This equation recognizes that the first transition from state i can be to either state j or to some other state k. If it is to state j, the first passage time is 1. Given that the first transition is to some state $k(k \neq j)$ instead, which occurs with probability p_{ik}, the conditional expected first passage time from state i to state j is $1 + \mu_{kj}$. Combining these facts, and summing over all the possibilities for the first transition, leads directly to this equation. **(تاج وسرحان ,2007)**

For the case of μ_{ij} where $j = i$, μ_{ii} is the expected number of transitions until the process returns to the initial state i, and is called the expected recurrence time for stat i.

After obtaining the steady-state probabilities $(\pi_0, \pi_1, \dots, \pi_m)$, these expected recurrence times can be calculated immediately as:

$$\mu_{ii} = \frac{1}{\pi_i} \ , \quad for \ \ i = 0,1, \dots, m \tag{11}$$

3. Analysis of Internal Migration Data in Egypt

In this section, we use the census data to find the transition probability matrix for both males, and females.

3.1. For Male Migrants

The first case is to find the transition probability matrix for male migrants between regions, where the numbers of people staying in the governorates of each region are neglected from the following table;

Table 1: The number of male migrants who migrate between regions of Egypt, in 2017.

Regions	Cairo region	Alexandria region	Delta region	Canal region	North Upper Egypt region	Central Upper Egypt region	Southern Upper Egypt region	Total
Cairo region	179822	6826	9749	9989	7480	3082	6603	223551
Alexandria region	27468	12230	4821	2839	847	484	1774	50463
Delta region	32991	9979	8876	10855	931	2584	2988	69204
Canal region	47345	5398	6153	17347	2092	970	3617	82922
North Upper Egypt region	33577	1907	1003	3500	763	291	1817	42858
Central Upper Egypt region	10085	1853	834	1094	674	727	1087	16354
Southern Upper Egypt region	13808	4605	1051	2720	740	1119	9100	33143
Total	345096	42798	32487	48344	13527	9257	26986	518495

Source: Central Agency for Public Mobilization and Statistics (CAPMAC), Egypt, 2017.

Based on Table (1) and the application of the maximum likelihood estimation by dividing each element in Table (1) by the sum of the corresponding row, then the transition probability matrix for male migration in Egypt in 2017 with state S= {C,A,D,CL,NU,CU,SU), where C stands for Cairo region, A stands for Alexandria region, D stands for Delta region, CL stands for Canal region, NU stands for North Upper Egypt region, CU stands for Central Upper Egypt region, SU stands for Southern Upper Egypt region is given as below;

$$
P = \begin{array}{c} \\ C \\ A \\ D \\ CL \\ NU \\ CU \\ SU \end{array}
\begin{array}{c}
\begin{array}{ccccccc} C & A & D & CL & NU & CU & SU \end{array} \\
\begin{bmatrix}
0.8044 & 0.0305 & 0.0436 & 0.0447 & 0.0335 & 0.0138 & 0.0295 \\
0.5443 & 0.2424 & 0.0955 & 0.0563 & 0.0168 & 0.0096 & 0.0351 \\
0.4767 & 0.1442 & 0.1283 & 0.1568 & 0.0135 & 0.0373 & 0.0432 \\
0.5710 & 0.0651 & 0.0742 & 0.2092 & 0.0252 & 0.0117 & 0.0436 \\
0.7834 & 0.0445 & 0.0234 & 0.0817 & 0.0178 & 0.0068 & 0.0424 \\
0.6166 & 0.1133 & 0.0510 & 0.0670 & 0.0412 & 0.0444 & 0.0665 \\
0.4166 & 0.1389 & 0.0317 & 0.0821 & 0.0223 & 0.0338 & 0.2746
\end{bmatrix}
\end{array} \qquad (12)
$$

We note from the matrix in (12) the main diagonal of the matrix represents the probabilities of people migrating between the governorates of the same region, while the rest of the probabilities represent migration between regions and each other.

Through the transition probability matrix, we find that the Cairo region is one of the most regions that retains male migrants, as it retains 80.44% of the total males within the region, while 19.56% of the total males migrate to the rest of the regions of the Republic. We also note that it is one of the most regions that attracts migrants, especially from Northern and Central Upper Egypt with the following proportions, respectively (78.34% and 61.66%).

We also find that the North Upper Egypt region is one of the least regions that retains male migrants, as it retains 1.78% of the total males within the region, while 98.22% of the total males migrate to the rest of the regions of the Republic.

To predict the migration probabilities for each region, we need both the transition probability matrix and the initial probability vector for migrating males between the republic's regions.

To find the initial probability vector, divide the sum of each row by the sum of all rows from

Table 1 as follows;

The initial probability vector for male migrants is as follows;

$$\begin{array}{cccccccc} & C & A & D & CL & NU & CU & SU \\ \pi_{2017}^{(0)} = (& 0.4312 & 0.0973 & 0.1335 & 0.1599 & 0.0827 & 0.0315 & 0.0639 \,) \end{array} \quad (13)$$

Vector (13) shows the distribution of male migrants between the regions of the Republic at the beginning of 2017 before the start of the transition process, where, the value 0.4312 represents the probability of male migrants from all regions to Cairo region at the initial time, while the value 0.0315 represents the probability of male migrants from all regions to Central Upper Egypt region and so on.

Through this vector and the transition probability matrix, we can find $\pi_{2018}^{(1)}$ through this equation $\pi^{(n)} = \pi^{(0)} * P^{(n)}$ as follows;

$$\begin{array}{cccccccc} & C & A & D & CL & NU & CU & SU \\ \pi_{2018}^{(1)} = (& 0.6656 & 0.0825 & 0.0627 & 0.0932 & 0.0261 & 0.0179 & 0.0520 \,) \end{array} \quad (14)$$

Through vector (14) and the total number of males for the year 2017, we found that the size of male migrants to Cairo region (345096) migrants, the size of male migrants to Southern Upper Egypt region (26986) migrants, and so on, where these numbers represent the size of in-migration for each region of the Republic, according to the 2017 census. To predict the number of male migrants for the coming period, we use the probability distribution $\pi^{(2)}$ from Table (2) and the total number of males for the year 2017, so we find that the size of male migrants to Cairo region (371535) migrants, the size of male migrants to Southern Upper Egypt region (23790) migrants, thus, the number of migrants can be predicted for each region of the Republic according to the same pattern.

Through the matrix of transition probability and the vector of initial probability, the migration probabilities for each region can be predicted for the coming periods, as the results were obtained using the R program in the following table:

Table 2: Initial state probabilities (bold case) and projected distribution for male migrants in Egypt.

States / Years	Cairo region	Alexandria region	Delta region	Canal region	North Upper Egypt region	Central Upper Egypt region	Southern Upper Egypt region
2017	**0.4312**	**0.0973**	**0.1335**	**0.1599**	**0.0827**	**0.0315**	**0.0639**
2018	0.6656	0.0825	0.0627	0.0932	0.0261	0.0179	0.0520
2019	0.7166	0.0658	0.0550	0.0713	0.0292	0.0161	0.0459
2020	0.7312	0.0599	0.0528	0.0665	0.0299	0.0159	0.0438
2021	0.7354	0.0580	0.0522	0.0654	0.0301	0.0158	0.0432
2022	0.7366	0.0574	0.0520	0.0651	0.0301	0.0158	0.0430
2023	0.7370	0.0572	0.0520	0.0650	0.0301	0.0158	0.0430
2024	0.7371	0.0572	0.0520	0.0650	0.0301	0.0158	0.0430
2025	0.7371	0.0572	0.0520	0.0649	0.0301	0.0158	0.0429
2026	0.7371	0.0572	0.0520	0.0649	0.0301	0.0158	0.0429
2027	0.7371	0.0571	0.0520	0.0649	0.0301	0.0158	0.0429
2028	0.7371	0.0571	0.0520	0.0649	0.0301	0.0158	0.0429
2029	0.7371	0.0571	0.0520	0.0649	0.0301	0.0158	0.0429

The results presented in Table (2) show that the percentage of male migrants to the Cairo region increases through the forecasting process for long periods, while the distribution of migrants to the rest of the regions decreases with time. We find that the number of male migrants to Cairo region in 2017 amounted to (345096) migrants, while we expect that the number of male migrants to Cairo region in 2029 will reach (382185) migrants, the size of male migrants to Southern Upper Egypt region in 2017 amounted to (26986) migrants, while we expect that the number of male migrants to Southern Upper Egypt region in 2029 will reach (22267) migrants.

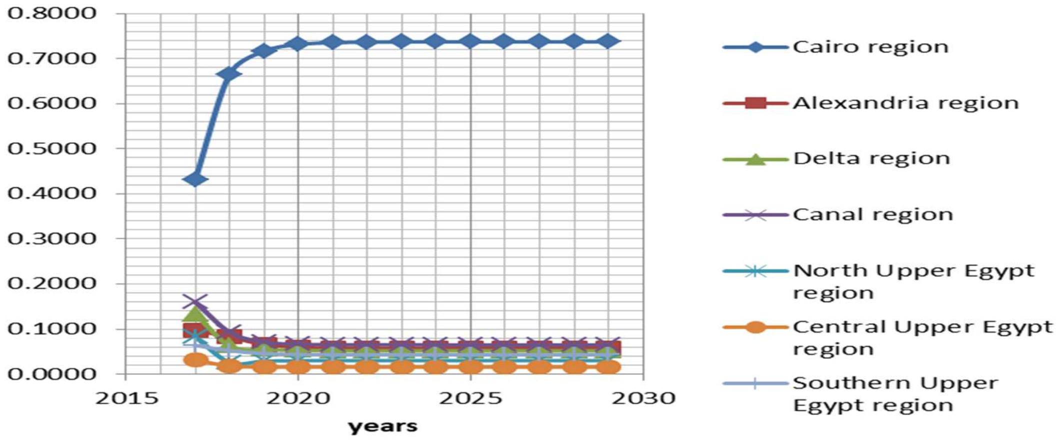

Figure 1: Initial and projected distribution for male migrants in Egypt.

Figure (1) shows the initial and stationary probabilities of male migrants which are computed from Markov chain model results in Table (2).

We also notice that the state of equilibrium was quickly reached after 20 periods of time, and the steady state probability vector is as follows:

$$\begin{array}{cccccccc} & C & A & D & CL & NU & CU & SU \\ \pi^{(20)}= & (0.73710 & 0.05715 & 0.05195 & 0.06494 & 0.03013 & 0.01578 & 0.04294) \end{array} \quad (15)$$

This probability vector represents equilibrium state for the migration process for each region of Egypt.

The steady state vector of male migrants can also be obtained through the equilibrium matrix as follows:

$$P^{(20)} = \begin{array}{c} \\ C \\ A \\ D \\ CL \\ NU \\ CU \\ SU \end{array} \begin{array}{c} \begin{array}{ccccccc} C & A & D & CL & NU & CU & SU \end{array} \\ \begin{bmatrix} 0.73710 & 0.05714 & 0.05195 & 0.06494 & 0.03013 & 0.01577 & 0.04294 \\ 0.73710 & 0.05714 & 0.05195 & 0.06494 & 0.03013 & 0.01577 & 0.04294 \\ 0.73710 & 0.05714 & 0.05195 & 0.06494 & 0.03013 & 0.01577 & 0.04294 \\ 0.73710 & 0.05714 & 0.05195 & 0.06494 & 0.03013 & 0.01577 & 0.04294 \\ 0.73710 & 0.05714 & 0.05195 & 0.06494 & 0.03013 & 0.01577 & 0.04294 \\ 0.73710 & 0.05714 & 0.05195 & 0.06494 & 0.03013 & 0.01577 & 0.04294 \\ 0.73710 & 0.05714 & 0.05195 & 0.06494 & 0.03013 & 0.01577 & 0.04294 \end{bmatrix} \end{array} \quad (16)$$

One of the important measures that can be calculated from the transition probability matrix and the steady state probability vector is the mean first passage time m_{ij} and the mean recurrence time m_{ii}, where the R program was used to obtain these averages , We denote the mean first passage time matrix by M where $M = m_{ij}$ as follow;

$$
M = \begin{array}{c} \\ C \\ A \\ D \\ CL \\ NU \\ CU \\ SU \end{array}
\begin{array}{c}
\begin{array}{ccccccc} C & A & D & CL & NU & CU & SU \end{array} \\
\begin{bmatrix}
0.00000 & 22.42383 & 21.32498 & 18.54105 & 32.65462 & 65.44471 & 30.92577 \\
1.854423 & 0.0000 & 19.90239 & 18.12399 & 33.41424 & 65.66220 & 30.67444 \\
1.977201 & 19.40948 & 0.00000 & 16.20497 & 33.51615 & 63.76575 & 30.34199 \\
1.790961 & 21.33789 & 20.48334 & 0.00000 & 33.05247 & 65.52372 & 30.37908 \\
1.402651 & 22.10793 & 21.70184 & 17.89793 & 0.00000 & 65.91368 & 30.52965 \\
1.707196 & 20.33295 & 21.02724 & 18.01987 & 32.49859 & 0.00000 & 29.70375 \\
2.145405 & 19.14688 & 21.40168 & 17.59514 & 33.25017 & 63.76550 & 0.00000
\end{bmatrix}
\end{array} \quad (17)
$$

Each element of this matrix represents the average number of transitions before migrating for the first time to a particular region. The value 1.854423 means that the first transition from Alexandria region to Cairo region takes place after two transitions or after two time periods.

The mean recurrence time vector is as follows:

$$
\begin{array}{cccccccc}
 & C & A & D & CL & NU & CU & SU \\
m_{ii} = (& 1.3567 & 17.49889 & 19.2492 & 15.3982 & 33.1840 & 63.391 & 23.2851 \;)
\end{array} \quad (18)
$$

The value 33.1840 means that the expected number of transitions until

the person returns to migrate to the North Upper Egypt region for the first time after his departure is 33 time periods.

3.2. For Female Migrants

The second case is to find the transition probability matrix for female migrants between regions, from the following table:

Table 3: The number of females migrants who migrate between regions of Egypt, in 2017.

Regions	Cairo region	Alexandria region	Delta region	Canal region	North Upper Egypt region	Central Upper Egypt region	Southern Upper Egypt region	Total
Cairo region	197679	6937	13924	11136	10710	3505	8223	252114
Alexandria region	29520	16034	7860	3375	1316	806	2344	61255
Delta region	35479	13234	15422	12016	1302	2498	2872	82823
Canal region	48178	6004	8676	19723	2193	1311	4089	90174
North Upper Egypt region	32210	1717	933	3406	1968	957	2103	43294
Central Upper Egypt region	10143	1768	639	917	1185	779	1569	17000
Southern Upper Egypt region	14108	4813	1252	2795	931	1662	13850	39411
Total	367317	50507	48706	53368	19605	11518	35050	586071

Source: (CAPMAC), Egypt, 2017

Based on, Table (3) the transition probability matrix is constructed as follows:

$$P = \begin{array}{c} \\ C \\ A \\ D \\ CL \\ NU \\ CU \\ SU \end{array} \begin{array}{c} \begin{array}{ccccccc} C & A & D & CL & NU & CU & SU \end{array} \\ \begin{bmatrix} 0.7841 & 0.0275 & 0.0552 & 0.0442 & 0.0425 & 0.0139 & 0.0326 \\ 0.4819 & 0.2617 & 0.1283 & 0.0551 & 0.0215 & 0.0132 & 0.0383 \\ 0.4283 & 0.1598 & 0.1862 & 0.1451 & 0.0157 & 0.0302 & 0.0347 \\ 0.5344 & 0.0666 & 0.0962 & 0.2187 & 0.0243 & 0.0145 & 0.0453 \\ 0.7439 & 0.0397 & 0.0216 & 0.0787 & 0.0455 & 0.0221 & 0.0485 \\ 0.5967 & 0.1040 & 0.0376 & 0.0539 & 0.0697 & 0.0458 & 0.0923 \\ 0.3580 & 0.1221 & 0.0318 & 0.0709 & 0.0236 & 0.0422 & 0.3514 \end{bmatrix} \end{array} \qquad (19)$$

Looking at the probability matrix, we find that Cairo region retains 78.41% of female migrants, and this percentage is slightly lower than the retention rate of male migrants. We also notice an increase in the percentage of retention of female migrants in the rest of the regions when compared to the previous probability matrix for male migrants.

The Cairo region remains the most attractive region for female migrants, and the North Upper Egypt region remains the most expelling region for female migrants.

The initial probability vector for female migrants is as follows:

$$\begin{array}{ccccccc} C & A & D & CL & NU & CU & SU \end{array}$$

$$\pi^{(0)} = (0.4302 \quad 0.1045 \quad 0.1413 \quad 0.1539 \quad 0.0739 \quad 0.0290 \quad 0.0672) \qquad (20)$$

The vector of $\pi_{2018}^{(1)}$ is as follows;

$$\begin{array}{ccccccc} C & A & D & CL & NU & CU & SU \end{array}$$

$$\pi_{2018}^{(1)} = (0.6268 \quad 0.0862 \quad 0.0831 \quad 0.0911 \quad 0.0335 \quad 0.0197 \quad 0.0598) \qquad (21)$$

As we know that the size of female migrants to Cairo region (367317) migrants, the size of female migrants to Delta region (48706) migrants, according to the 2017 census. Now we predict the number of female migrants for the coming period, we use the probability distribution $\pi^{(2)}$ from Table (4) and the total number of females for the year 2017, we find that the size of female migrants to Cairo region (395748) migrants, the size of female migrants to Delta region (42929) migrants, and so on.

The prediction results for the distribution of female migrants are displayed in Table (4):

Table 4: Initial state probabilities (bold case) and projected distribution for female migrants in Egypt.

States / Years	Cairo region	Alexandria region	Delta region	Canal region	North Upper Egypt region	Central Upper Egypt region	Southern Upper Egypt region
2017	**0.4302**	**0.1045**	**0.1413**	**0.1539**	**0.0739**	**0.0290**	**0.0672**
2018	0.6268	0.0862	0.0831	0.0911	0.0335	0.0197	0.0598
2019	0.6753	0.0698	0.0732	0.0724	0.0363	0.0178	0.0552
2020	0.6956	0.0611	0.0690	0.0667	0.0376	0.0174	0.0526
2021	0.6972	0.0603	0.0686	0.0663	0.0377	0.0174	0.0523
2022	0.6978	0.0600	0.0685	0.0662	0.0378	0.0174	0.0523
2023	0.6980	0.0599	0.0685	0.0662	0.0378	0.0174	0.0522
2024	0.6981	0.0599	0.0685	0.0662	0.0378	0.0174	0.0522
2025	0.6981	0.0599	0.0685	0.0662	0.0378	0.0174	0.0522
2026	0.6981	0.0599	0.0685	0.0662	0.0378	0.0174	0.0522
2027	0.6981	0.0599	0.0685	0.0662	0.0378	0.0174	0.0522
2028	0.6981	0.0599	0.0685	0.0662	0.0378	0.0174	0.0522
2029	0.6981	0.0599	0.0685	0.0662	0.0378	0.0174	0.0522

The results presented in Table (4) show that the percentage of female migrants to Cairo region increases through the forecasting process for long periods, while the distribution of migrants to the rest of the regions decreases with time. We find that the number of female migrants to Cairo region in 2017 amounted to (367317) migrants, while we expect that the number of female migrants to Cairo region in 2029 will reach (409135) migrants, the size of female migrants to Delta region in 2017 amounted to (48706) migrants, while we expect that the number of female migrants to Southern Upper Egypt region in 2029 will reach (40125) migrants.

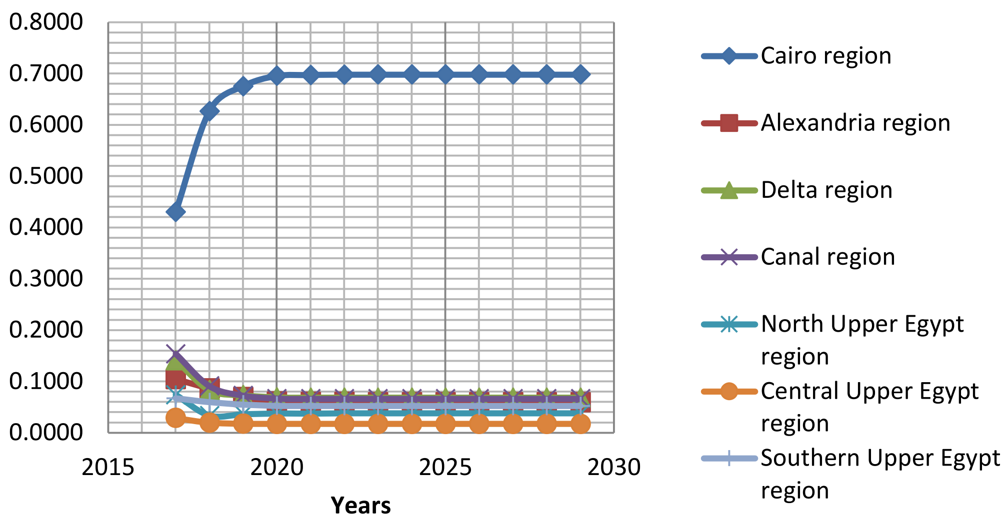

Figure 2: Initial and projected distribution for female migrants in Egypt.

Figure (2) shows the initial and stationary probabilities of female migrants which are computed from Markov chain model results in Table (4).

The steady state probability vector is as follows:

$$\begin{array}{ccccccc} C & A & D & CL & NU & CU & SU \end{array}$$
$$\pi^{(20)} = (0.68809 \quad 0.05990 \quad 0.06846 \quad 0.06617 \quad 0.03780 \quad 0.01735 \quad 0.05220) \qquad (22)$$

The equilibrium matrix is as follows:

$$P^{(20)} = \begin{array}{c} \\ C \\ A \\ D \\ CL \\ NU \\ CU \\ SU \end{array} \begin{array}{c} \begin{array}{ccccccc} C & A & D & CL & NU & CU & SU \end{array} \\ \begin{bmatrix} 0.68809 & 0.05990 & 0.06846 & 0.06617 & 0.03780 & 0.01735 & 0.05220 \\ 0.68809 & 0.05990 & 0.06846 & 0.06617 & 0.03780 & 0.01735 & 0.05220 \\ 0.68809 & 0.05990 & 0.06846 & 0.06617 & 0.03780 & 0.01735 & 0.05220 \\ 0.68809 & 0.05990 & 0.06846 & 0.06617 & 0.03780 & 0.01735 & 0.05220 \\ 0.68809 & 0.05990 & 0.06846 & 0.06617 & 0.03780 & 0.01735 & 0.05220 \\ 0.68809 & 0.05990 & 0.06846 & 0.06617 & 0.03780 & 0.01735 & 0.05220 \\ 0.68809 & 0.05990 & 0.06846 & 0.06617 & 0.03780 & 0.01735 & 0.05220 \end{bmatrix} \end{array} \qquad (23)$$

The mean recurrence time vector is as follows:

$$\begin{array}{ccccccc} C & A & D & CL & NU & CU & SU \end{array}$$
$$m_{ii} = (1.43246 \quad 16.69404 \quad 14.60632 \quad 15.11149 \quad 29.45389 \quad 57.62111 \quad 19.15487) \qquad (24)$$

The value 16.69404 means that the expected number of time periods or transitions until the person returns to migrate to Alexandria region for the first time after his departure is 17 time periods.

The matrix of the mean first passage time is as follows:

$$
\begin{array}{c|ccccccc}
 & C & A & D & CL & NU & CU & SU \\
\hline
C & 0.00000 & 22.17122 & 17.18268 & 18.43511 & 26.52587 & 59.66962 & 28.22149 \\
A & 2.081834 & 0.0000 & 15.51668 & 17.97376 & 27.36670 & 59.60954 & 28.00344 \\
D & 2.184878 & 18.61333 & 0.0000 & 16.23611 & 27.54171 & 58.53567 & 28.05003 \\
CL & 1.947763 & 20.90136 & 16.24362 & 0.0000 & 27.21310 & 59.54050 & 27.77769 \\
NU & 1.509655 & 21.91118 & 17.71561 & 17.85242 & 0.0000 & 59.17296 & 27.74459 \\
CU & 1.816039 & 20.27406 & 17.38846 & 18.19408 & 25.88591 & 0.0000 & 26.44919 \\
SU & 2.426815 & 19.09309 & 17.49852 & 17.73016 & 27.34406 & 57.23001 & 0.0000
\end{array}
\quad = M \qquad (25)
$$

The value of 2.081834 means that the expected number of transitions for the migrant person before his first arrival to Cairo region is two transitions, given that he is currently in Alexandria region, and this means that the first transition from Alexandria region to Cairo region takes place after two steps or after two time periods.

From the above, we conclude that Cairo region is the most attractive region for internal migration streams, whether for males, females, while the rest of the regions expel people, and the rate of expulsion varies in each of them and the flow of migration streams are as follows: 1) from Cairo region to Delta and Canal region, 2) from Alexandria region to Cairo and Delta region, 3) from Delta region to Cairo, Canal, and Alexandria region, 4) from Canal region to Cairo, Delta, and Alexandria region, 5) from North Upper Egypt region to Cairo and Canal region, 6) and from Central and Southern Upper Egypt to Cairo and Alexandria region.

4. Conclusion

The goal of achieving population redistribution remains at the forefront of the population goals in Egypt, and internal migration is one of the most important mechanisms for implementing this goal by directing migration flows from areas with high population density to areas with less density. Egypt witnessed internal population movements during the last census period, as our study aimed to analyze these movements between the regions of Egypt using the Markov chain model to predict the numbers of migrants for the coming periods in an attempt to help the government in setting development plans and distributing investments to the regions of the Republic.

In this study, we found that Cairo region is the main destination for most migrants, whether from within the region or from other regions, whether male or female, as it recorded the highest rate of attraction for internal migration streams, attracting 66.6% of total female migrants and 62.7% of total male migrants, whether from within the region or from other regions, while the rest of the regions of the Republic are population-expelling regions, and the rate of expulsion varies in both of them, we also found that the North Upper Egypt region is the most region that expels the migrants, especially to Cairo region, and it is also the least region that retains migrants, whether male or female. It is clear from the forecasting process that the probabilities of migration for each region will be stable in Egypt after 20 years of prediction and the distribution of male or female migrants to Cairo region increases through the forecasting process for long periods, while the distribution of migrants to the rest of the regions decreases with time. It is necessary to limit the migration flows coming to Cairo region, direct the population to other regions, and investigate the reasons for not attracting these regions to migrants.

References

Agbinya, J. I. (2020*). Applied data analytics-principles and applications*. River Publishers, Denmark.

Ching, W. K. & Ng, M. K. (2006). *Markov chains: models algorithms and applications*. Springer Science & Business Media, Hong Kong.

Collins, L. (1975). *An introduction to Markov chain analysis*. Geo Abstracts & University of East Anglia, Great Britain.

Compton, P. A. (1968). Internal migration and population change in Hungary between 1959 and1965. *Transactions of the Institute of British Geographers*, 111-130.

Eledum, H. Y., & Hussein, Y. (2018). *Lecture notes on stochastic processes*. University of Tabuk.

Garrocho, C., Lopez, E. J., & Lobato, J. (2016). Modeling interstate migration in Mexico: Static Markov chains versus dynamic Markov chains with moving averages. *Population Papers*, 22(90), 109-144.

Gulshan, J., Mina, M. M., & Hossain, S. S. (2015). Migration pattern in Bangladesh: a covariate-dependent Markov model. *Journal of Applied Statistics*, 42(7), 1519-1530.

Hierro, M. (2008). Modelling the dynamics of internal migration flows in Spain. *Papers in Regional Science*, 88(3), 683-692.

Medhi, J. (2003). *Stochastic models in queueing theory* (2nd ed.). Elsevier Science, USA.

Oliver, C. (2009). *Markov processes for stochastic modeling*. Elsevier Academic press, USA .

Olofsson, P., &Andersson, M. (2010). *Probability, statistics, and stochastic processes* (2nd ed.), John Wiley & Sons, New Jersey.

Sericola, B. (2013). *Markov chain: theory, algorithms and applications*. Great John Wiley &Sons, United Kingdom.

Stewart, W. J. (2009*). Probability, Markov chain, queues, and simulation*. Princeton University Press, United Kingdom.

Todaro, M.P. (1976). *Internal Migration in Developing Countries: A Review of Theory, Evidence, Methodology and Research, Priorities*. International Labour Office, Geneva.

Usman, M., Elfaki, F., Barusman, Y., &Widiarti, W. (2015). Markov chain analysis and mechanism of migration in Indonesia in the period 1980-2010. *Journal of Engineering and Applied Sciences*, 10(22), 17256-17264.

Venkatesan, G., & Sasikala,V. (2018).Statistical analysis on migrants using Markov chain model. *International Journal of Creative Research Thoughts*, 6(1), 860-866.

Yip, p., Soleymain, M., Wat, K. P., Pinkney, E., &Lam, K.F. (2020).Modeling internal movement of children Born in Hong Kong to nonlocal mothers. *International Journal of Environmental Research and Public Health*, DOI 10.3390/ijerph 17155476, 1-12.

المراجع العربية

الجهاز المركزي للتعبئة العامة والإحصاء. (2019). النتائج النهائية لتعداد السكان والظروف السكنية بالتعداد العامل للسكان والإسكان والمنشآت لعام 2017 نتائج العينة (الاستمارة المطولة) جدول 1-21.

ازهري, نور. (2017). استخدام طوريات ماركوف المخفيه في التعرف علي الصور والرموز [رسالة ماجستير منشوره, جامعة تشرين].

الحنفي, يحي عبد الرحمن. (2018). التنبؤ باحتمالات تغير القروض الزراعيه باستخدام نموذج سلاسل ماركوف. *معهد البحوث الزراعيه-الجيزه*,3(96),1229-1257.

تاج, لطفي وسرحان,عمار محمود علي. (2007). *مقدمه في العمليات العشوائيه* .النشر العلمي والمطابع, جامعه الملك سعود. الرياض.

حسانين، محمد أحمد علي. (2009). الهجرة الداخلية في مصر خلال الفترة (1960-1996): دراسة جغرافية [رسالة دكتوراة غير منشورة]. كلية الآداب, جامعة القاهرة.