A Markov Model to Study Covid-19 Herd Immunization

Jorge Luis Romeu, Ph.D.
https://www.researchgate.net/profile/Jorge_Romeu
http://web.cortland.edu/romeu/; Email: romeu@cortland.edu

1.0 Introduction

This second Markov model paper studies *Covid-19 Herd Immunization approach* which assumes that the virus will infect a large percentage of the population, thus preventing further community spread. In our previous paper *A Markov Chain Model for Covid-19 Survival Analysis* (found in: https://www.researchgate.net/publication/343021113_A_Markov_Chain_Model_for_Covid-19_Survival_Analysis) we assumed that there was neither a vaccine nor treatment for *Covid-1,* and that if current infection rates remained unchecked, everyone would eventually die. In the present paper we assume that Covid-19 survivors become immune, thence, cannot become re-infected again, as occurred in our first Markov Chain paper. This creates the *Herd Immunity*.

There is much debate about employing *Herd Immunity* as an alternative solution for Covid-19. We implement a Markov Chain model to quantitatively analyze such situation. With the Markov model, *we will obtain* (1) **the probability of death or immunization** of a Patient; and **(2) their expected times to death** (or immunization), when starting from the different states in the Space.

We have previously written *An Example of Survival Analysis Applied to Covid-19 Data*, found in https://www.researchgate.net/publication/342583500_An_Example_of_Survival_Analysis_Data_Applied_to_Covid-19, then *Multivariate Statistics in the Analysis of Covid-19 Data*, and *More on Applying Multivariate Statistics to Covid-19 Data*, both of which can also be found in: https://www.researchgate.net/publication/341385856_Multivariate_Stats_PC_Discrimination_in_the_Analysis_of_Covid-19 and, as all the already cited, also in our *ResearchGate* web page: https://www.researchgate.net/publication/342154667_More_on_Applying_Principal_Components_Discrimination_Analysis_to_Covid-19 These latter statistical methods provide useful tools for classification of states, regions, counties etc., according to levels of infection and other metrics.

We have also written a tutorial on the use of *Design of Experiments Applied to the Assessment of Covid-19*. It provides an example of the use of DOE for assessing and controlling the appropriate levels of infection in our states and regions. It can also be found in our *ResearchGate* web page: https://www.researchgate.net/publication/341532612_Example_of_a_DOE_Application_to_Coronavarius_Data_Analysis We have written *an evaluation of* the results of 25 years *off-shoring tens of thousands American jobs*, and *its impact on US preparedness* to fight the Coronavarus Pandemic. It can also be found in: https://www.researchgate.net/publication/341685776_Off-Shoring_Taxpayers_and_the_Coronavarus_Pandemic And we have written a short study on the use of *reliability methods in the design and operation of ICU* units, that can also be found in:

https://www.researchgate.net/publication/342449617_Example_of_the_Design_and_Operation_of_an_ICU_using_Reliability_Principles

All above work is part of our *pro-bono collaboration to the American struggle against Covid-19,* based on our *Proposal for Fighting Covid-19 and its Economic Fallout* that can be read in: https://www.researchgate.net/publication/341282217_A_Proposal_for_Fighting_Covid-19_and_its_Economic_Fallout Such proposal encourages retired professionals like this author to provide *pro-bono* analyses, based on our long work experience.

2.0 A Markov Chain Model with Two Absorbing States.

In this paper we reanalyze the Covid-19 situation using a more complex model. The new Markov Chain is defined over a *larger (eight element) state space,* yielding a more realistic model, one which is more flexible and allows for more options. This Markov Chain models a *non-recurrent process* that, moving through a number of *Transient states,* eventually *leads us to one of two Absorbing States: Death or Immunization*, where patients will remain forever. These new model features will introduce several technical differences.

Because the *state space* contains both transient and recurring (absorbing) states there is no steady state solution. Instead, we obtain (1) the *long run probabilities of dying or becoming immunized*, as well as (2) *their expected times*, both of these when starting in any of the Transient states.

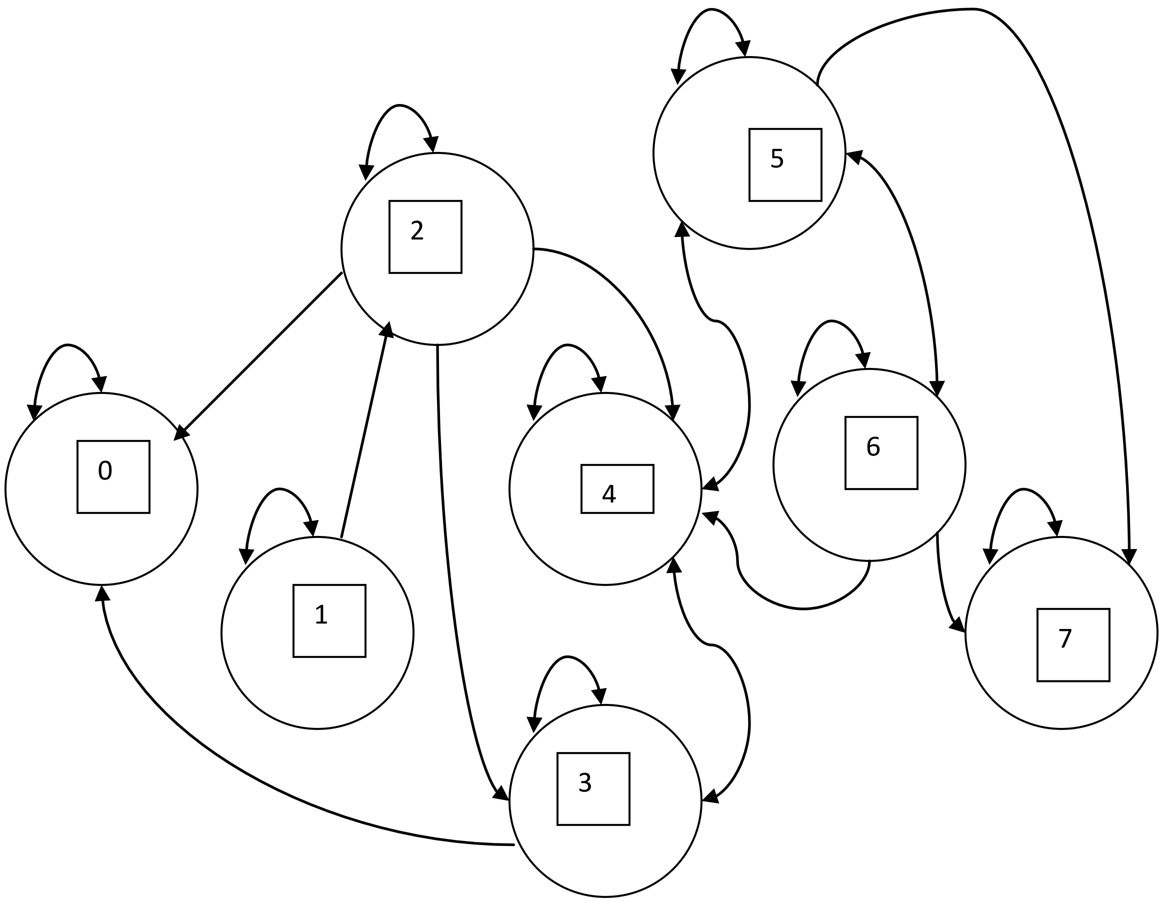
Let $\{X_n; n >= 0\}$ be a *Markov Chain* over an *eight-element State Space* defined as:

(0) Covid-Immunized population (an absorbing state);
(1) *Non Infected* persons in the General Population;
(2) *Infected persons* (but asymptomatic; i.e. not known to be such);
(3) *Infected persons Detected and Isolated*; (after symptoms, or Covid-19 tests positive)
(4) *Hospitalized patients* (after becoming ill with Colvid-19);
(5) *Patients in the ICU* (very sick);
(6) Patients in a Ventilator (critical);
(7) *Death* (an absorbing state)

The *Transition Probability Matrix P* for this *Markov Chain* is:

State	Immune	UnInfected	Infected	Isolated	Hospital	ICU	Ventilator	Dead
0	1.0	0.00	0.00	0.0	0.0	0.0	0.0	0.0
1	0.0	0.96	0.04	0.0	0.0	0.0	0.0	0.0
2	0.2	0.00	0.40	0.2	0.2	0.0	0.0	0.0
3	0.2	0.00	0.00	0.6	0.2	0.0	0.0	0.0
4	0.0	0.00	0.00	0.1	0.8	0.1	0.0	0.0
5	0.0	0.00	0.00	0.0	0.3	0.3	0.3	0.1
6	0.0	0.00	0.00	0.0	0.1	0.3	0.3	0.3
7	0.0	0.00	0.00	0.0	0.0	0.0	0.0	1.0

<u>Figure 1: The *State Space Diagram* for this Markov Chain is:</u>



The Markov Chain *unit time* is *a day*. *Transitions* refer to the State changes that occur from *one morning to the following morning*. State (0) corresponds to the *Immunized* population. State (1) corresponds to persons not yet infected with Covis-19. The daily rate of infection is 4% (which means that 96% remain uninfected). State (2) corresponds to persons that *contracted Covid-19*, but are *not identified as infected*. They spread the virus in the community (unless community use of face covering, social distancing and other preventive provisions are observed). Infected are asymptomatic and not aware of their infection. Some then develop symptoms, are tested, and are quarantined or isolated at home. If they become very ill, they are sent to the hospital. State (3) is *Isolation at home*. Persons either improve and become Immune, or worsen and are then sent to the hospital. State (4) corresponds to *hospital patients*, receiving Covid-19 treatment in *wards*. They can recover and are sent home, or get worse, in which case they are placed in ICUs. State (5) corresponds to *patients in ICU*. Of these, 30% improve and are sent back to the Ward; 30% are placed on Ventilators, and 10% die. About 30% of those are placed in *Ventilators* (State (6)), improve, and are returned to the ICU or to the ward. Finally, 15% *die (the Absorbing State (7)).*

We obtain, from P, the Sub-Matrix Q, corresponding to the six Transient states:

Matrix Q of the Six Transient States:

UnInf	Infec	Isol	Hosp	ICU	Ventilator
0.96	0.04	0.0	0.0	0.0	0.0
0.00	0.40	0.2	0.2	0.0	0.0
0.00	0.00	0.6	0.2	0.0	0.0
0.00	0.00	0.1	0.8	0.1	0.0
0.00	0.00	0.0	0.3	0.3	0.3
0.00	0.00	0.0	0.1	0.3	0.3

We then subtract matrix Q from the Identity Matrix, yielding (I-Q),

UnInf	Infec	Isol	Hosp	ICU	Ventilator
0.04	-0.04	0.0	0.0	0.0	0.0
0.00	0.60	-0.2	-0.2	0.0	0.0
0.00	0.00	0.4	-0.2	0.0	0.0
0.00	0.00	-0.1	0.2	-0.1	0.0
0.00	0.00	0.0	-0.3	0.7	-0.3
0.00	0.00	0.0	-0.1	-0.3	0.7

We invert the above: **Matrix inverse S = $(I-Q)^{-1}$**

The Potential Matrix (of which S is a sub-matrix) contains the *Long-Run Sojourns* (the average number of visits) *to each of the States* (matrix columns), when starting from any other of the States, represented by the matrix rows. The Sojourns that occur outside S are either Zero (if the states cannot be reached) or Infinite (if it is to an Absorbing state).

State	UnInfect	Infected	Isolated	Hospital	ICU	Ventilator
1	25	1.66667	2.22222	5.5556	0.97222	0.41667
2	0	1.66667	2.22222	5.5556	0.97222	0.41667
3	0	0.00000	3.88889	5.5556	0.97222	0.41667
4	0	0.00000	2.77778	11.1111	1.94444	0.83333
5	0	0.00000	1.66667	6.6667	2.91667	1.25000
6	0	0.00000	1.11111	4.4444	1.52778	2.08333

Adding the matrix S rows, calculated for any row (State), yields Average time to Death.

For example, the *average number of days* an *uninfected* person *spends* before becoming infected is 25 days. The number of *days spent in the hospital*, is 11.1 days. The *number of days* a patient *spends in an ICU* unit is *2.92 days*. *By adding the rows*, we obtain *the average number of days* for a person, say initially hospitalized, to go through the ICU, then the Ventilator, until finally *passing away (dying)*. The average is: 11.11+1.94+0.83 = 13.9 days. We can do likewise with all other rows, and obtain the *average times to death*, starting from any one of the Transient states.

3.0 Results for the Two Absorbing States

From the previous section we can extract very useful results, related to patient probabilities of Absorption, and the corresponding average times to Absorption (of death or immunization).

Table #1: average times to death from each of the transient states

Starting State for any Individual	Average Time to Pass Away (Die)
From Time of Initial Infection (undetected)	=1.667+2.22+5.556+ 0.972+0.417=10.83 days
From the Time of Infection/Isolation	=3.889+5.556+0.972+0.4167=10.83 days
From the Time of Hospitalization	=11.11+1.94+0.83 = 13.9 days
From the Time of entering an ICU	=2.91667 + 1.25000 = 4.17 days
From the Time of entering a Ventilator	2.08 days

We now calculate the *probabilities of a patient dying, or becoming immunized, starting from any of the transient states*. We calculate these probabilities *using Matrix B*, a sub-matrix of transient states. B is constructed by reordering the rows and columns of P, in such way that all transient states are placed first, and all absorbing ones are placed last. Then, Matrix B corresponds to the rows and columns above the Identity sub-matrix, corresponding to the Absorbing states.

Matrix B:

```
0.0  0.0
0.0  0.2
0.0  0.2
0.0  0.0
0.1  0.0
0.3  0.0
```

Matrix G = SB; Prob. of Ever Reaching Death/Immunization:

Death	Inmunization
0.222222	0.777778
0.222222	0.777778
0.222222	0.777778
0.444444	0.555556
0.666667	0.333333
0.777778	0.222222

Matrix G provides the probabilities of ever reaching one of the two Absorbing States (death or immunization), from any of the Transient states.

Table #2 Probability of Dying or Becoming Immune, starting from a Transient State

Starting State	Probability of Dying	Probability Immunization
Uninfected	0.222222	0.777778
Infected/undetected	0.222222	0.777778
Infected/Isolated	0.222222	0.777778
Hospitalized	0.444444	0.555556
In ICU	0.666667	0.333333
On Ventilator	0.777778	0.222222

We see how the *Probability of Dying from Covid-19* is much smaller than the *Probability of Recovery and of becoming Immunized*, for the first four States (up to Infected, when detected as such by symptoms or tests, but displaying minor health issues, and in isolation). These two probabilities are much closer, when the patient is sick and hospitalized (0.44 v. 0.55). Finally, once in an ICU or a Ventilator, which indicates that the patient is very sick, the Probability of Dying is much larger than the Probability of Recovering and becoming Immune.

4.0 Mathematics

The key section of the *Potential Matrix R* is $(I-Q)^{-1} = S$, which provides the *Sojourns*, or average number of times that different Transient states are visited, if coming from other Transient states. Adding these up we obtain an *average times to Death*, when starting at any Transient state.

Matrix S in turn, provides the relevant part of *Matrix F*, which contains the *probabilities of ever reaching an absorption state (Death or Immunization*) starting from every Transient State. Other Matrix F Probabilities relating Transient and Recurrent states are defined as: (1) *Zero*, if between Recurrent and Transient states (it is *impossible to leave an absorbing state*) and (2) *Unit, the sure event* (if there is a single absorbing state), *or the rows of B add to Unit* (if there is more than one Absorbing State). This result implies that *every person will eventually arrive*, independently of their initial state, *to one of the Absorbing States*.

The probabilities of ever reaching a State, between two Transient states (i,j), which form part of *Hitting Matrix F,* are obtained from sub-matrix S of R, above mentioned, through the formulas:

$$F(i, j) = R(i,j) / R(j,j); \qquad \text{and} \qquad F(j,j) = 1 - 1 / R(j,j);$$

The above-mentioned *performance measures are used to calculate* many of the required health care facility *Logistics*. Statistics and mathematical *models* help *evaluate different alternatives,* and are used in what *Operations Researchers call the What If Game* (e.g. what would happen if infection rate were 9% instead of 4%?).

Further readings about Markov Processes can be found in the Bibliography Section.

5.0 Discussion

Results from the present Markov model, complement results obtained in our previous papers, as different models address different aspects of the Covid-19 problem. For example, the *survival model* in Romeu (2020) can be *used to evaluate the probability of dying* of a patient, given its co-morbidities, and can be used *in Triage situations*, when, say, Ventilators are less than patients.

The present model, due to its specific State Spaces E_j and its transition probabilities (X_{ij} for i,j = 0, … ,7) is particularly *useful to assess the Immunization theory*. For example, State (0) defines the Immunized population, and State (2) defines the infected but yet undetected one. Through the corresponding infection (transition) rates we can *study the impact of the different levels of these rates on the probabilities of dying or becoming immunized*. Differing infection rates depend on compliance with public health measures such as face covering, social distancing, etc. Thence, by assigning different values to these rates, their impact can then be assessed and compared.

Similarly, transition rates into and out of States (4), (5) and (6), related to hospitalization, depend on the load such hospitals and health care practitioners have, which in turn depend on infection rates, which in turn also depend on community compliance with public health measures such as face covering, social distancing, etc.. By assigning specific values to these, resulting patient care efficiency can be assessed, and subsequent Logistics problems can be addressed.

Probabilities of dying and immunization, from States E_j, combined with population estimates, can help determine the number of cases, of deaths, of immunized population, etc. that result from the different scenarios, which are in turn defined by the different transition probabilities or rates.

Statistics and mathematical models are tools that help answer such health questions, as well as to compare the performance of different alternatives, in a more objective, non-partisan way.

6.0 Conclusions

The *Markov Chain developed in this paper*, as well as its transition rates, and its transient and recurrent states, were all *based on our readings about Covid-19*. This author is not a public health specialist, but a statistician. *Our models were built using our 45+ years of statistical experience* in modeling and data analysis, in several research and academic organizations. *Our intent is to place them in the hands of health specialists*, so they can redo them with better information.

Results from Markov Chain models, most useful for Covid-19 problems, have to do with the probabilities and times to absorption (death and immunization) and are given in Section 3.0.

By *changing the transition rates* in the recurrent Markov Chain of Section 2.0, we can study how *they affect the absorption probabilities* (death or immunization) in each state, and *compare* the efficient and inefficient *models*. This may also help *establish an acceptable infection rate*. Times spent in a State (*Sojourn)* help *estimate* the *required size of the facilities* that will treat patients.

The results in Section 3.0 are more helpful. The probabilities of ever reaching every other state, and the Time spent (Sojourns) in them, may help *find logistic parameters* to ensure a sufficiently well-equipped health care facility, that can serve the public efficiently.

Combining results from the present *Markov Chain Model,* with those in our *previous Survival Model*, may *provide* doctors with more *neutral rules* to establish *Triage procedures*, may such procedures ever become necessary.

Finally, *our statistical models* are not intended to compete with, but to *complement and enrich*, the *models developed by Epidemiologists* and Public Health professionals. One example[1]:

The total population is partitioned into eight stages of disease: S, susceptible (uninfected); I, infected (asymptomatic or pauci-symptomatic infected, undetected); D, diagnosed (asymptomatic infected, detected); A, ailing (symptomatic infected, undetected); R, recognized (symptomatic infected, detected); T, threatened (infected with life-threatening symptoms, detected); H, healed (recovered); E, extinct (dead). The interactions among these stages are shown in Fig. 1. We omit the probability rate of becoming susceptible again after having recovered from the infection. Although anecdotal cases are found in the literature27, the re-infection rate value appears negligible.

The reader will recognize how their eight stages of the disease are somewhat similar to our Eight Element State Space, and how said article's *Figure 1* is also somewhat similar to our own *State Space Diagram*, presented in Section 2.0. Modeling tools used are definitely different.

Both models are fine and pursue a similar objective: help study and overcome Covid-19. We are allies, not competitors. As George Box once said: *all models are wrong; some models are useful*.

Let's work together to defeat Covid-19.

[1] See: https://www.nature.com/articles/s41591-020-0883-7#Sec2

Bibliography

Taylor, H. and S. Karlin. An Introduction to Stochastic Modeling. Academic Press. NY. 1993.

Cinlar, E. Introduction to Stochastic Processes. Prentice Hall. NJ. 1975.

Heyman, D. and M. Sobel. Handbooks in Operations Research and Management Science. *Vol. 2: Stochastic Models*. North Holland. Amsterdam. 1990.

Romeu, J. L. RAC START: Vol. 11, No. 6: *Availability*. https://web.cortland.edu/matresearch/AvailabilitySTART.pdf

Romeu, J. L. RAC Journal: First Quarter; 2000: *A Discussion of SW Reliability Modeling Problems*. https://web.cortland.edu/matresearch/SWRelArtRomeuRAC1ST_Q2000.pdf

Romeu, J. L. QR&CII Tutorial Series, Vol. 1, No. 2*: Infant Mortality in Reliability Studies for the Weibull*. https://web.cortland.edu/matresearch/InfantMortStudyQR&CII.pdf

Romeu, J. L (2020) *An Example of Survival Analysis Applied to Covid-19 Data*. Research Report https://www.researchgate.net/publication/342583500_An_Example_of_Survival_Analysis_Data_Applied_to_Covid-19

About the Author:

Jorge Luis Romeu retired Emeritus from the State University of New York (SUNY). He was then, for sixteen years, a Research Professor at Syracuse University, where he is currently an Adjunct Professor of Statistics. Romeu worked for many years as a Senior Research Engineer at the Reliability Analysis Center (RAC), an Air Force Information and Analysis Center operated by IIT Research Institute (IITRI). Romeu received seven Fulbright assignments: in Mexico (3), the Dominican Republic (2), Ecuador, and Colombia. He holds a doctorate in Statistics/O.R., is a C. Stat Fellow of the Royal Statistical Society (RSS), a Member of the American Statistical Association (ASA) and of the American Society for Quality (ASQ). Romeu is a Past ASQ Regional Director, and holds Reliability and Quality ASQ Certifications. He created and directs the *Juarez Lincoln Marti Int'l. Education Project* (https://web.cortland.edu/matresearch/) which is dedicated to support higher education in Ibero-America. Readers can find all our work on Coronavarus (as well as many other subjects) in our *Quality, Reliability and Continuous Improvement Institute* (QR&CII) web page:

https://web.cortland.edu/romeu/QR&CII.htm