Epidemic Spreading: Brief Notes

M.Cattani(mcattani@if.usp.br)
Instituto de Física da Universidade de São Paulo

Abstract

This paper was written to graduate and postgraduate students of Physics and Mathematics. Was done a brief analysis of the mathematical modeling of the epidemic spreading using stochastic and deterministic approaches. We have shown, taking into account the Master Equation and Fokker-Planck Equation, that the deterministic approach gives a good description of the epidemic time evolution.

***Key words*:** epidemic spreading; Master and Fokker-Planck equations; deterministic equations.

(I)Introduction.

Motivated by the current covid19 pandemic, this paper was written addressed to graduate and postgraduate students of Physics and Mathematics. Are briefly analyzed mathematical models proposed to describe the time evolution of the epidemic spreading. Only main aspects of these models have been discussed. Detailed analysis can be found in references mentioned in the text. Throughout human history, there have been a number of pandemics [1] of diseases like, for instance, Smallpox, Tuberculosis, Black Death, that killed an estimated 75-200 million people in the 14th century; 1918 Influenza (Spanish Flu) that killed an estimated 50-100 million people in the 19th century, Cholera and Infantile Paralysis. We have now the covid19 plague originated in China, in late December 2019.According to media reports, more than 200 countries and territories have been affected. Until 22july2020 the number of people infected had reach ~15,000,000 worldwide, of whom ~9,400,000 have recovered. The death toll is ~630.000. [1] In Brazil, we have ~2.200,000 infected,~100,000 deaths and ~1.500,000 recovered. In **Section 1** are seen general aspects about the mathematical modeling of infection diseases. In **Section 2** we present the stochastic approach. In **Section 3**, the epidemic outbreak. In **Section 4**, are presented examples of stochastic models . In **Section 5** are shown estimations done with stochastic and deterministic models and conclusions. Finally, in **Section 6** are presented some deterministic models.

(1)Mathematical modeling of epidemic spreading.

Mathematical models can project how in infectious diseases progress

to show the likely outcome of an epidemic and help inform public health interventions. Models use basic assumptions or collected statistics along with mathematics to find parameters for various infectious diseases and use those parameters to calculate the effects of different interventions necessary to help the infected population.[2-4]

The earliest account of mathematical modeling of spread of disease was carried out in 1760 by Daniel Bernoulli.[2] He created a mathematical model to defend the practice of inoculating against smallpox.

In the 20th century, William Hamer[2] and Ronal Ross[2] applied the *law of mass action* to explain the epidemic behavior. The 1920s saw the emergence of compartmental models. The Kermack-McKendrick[2] epidemic model(1927) and the Reed-Frost[2] epidemic model (1928) both describe the relationship between **Susceptible(S)**, **Infected(I)** and **Recovered(R)** individuals in a population. The predicted behavior of outbreaks are very similar to that observed in many recorded epidemics.[2]**.**These theoretical studies of the epidemic spreading[3,4] started with the employment of ordinary differential-equations of the first order in time, which became known as the **deterministic approach**. This approach however, does not describe, in a explicit manner ,the random fluctuations occurring in a real epidemic spreading. These fluctuations are seen in **Figure 1** which shows, for instance, the infected number **I**(t), as a function of time, due to the covid19 plague at São Paulo (**Brazil-2020**)

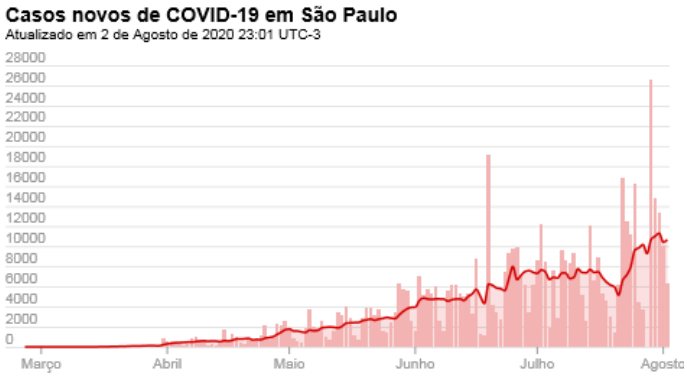

Figure 1.Time evolution of the infected number I(t) by the covid19at São Paulo.
From March up to August**.** Are clearly seen the random fluctuations in the **I**(t) number.

To take into account random fluctuations, Bartlett[2,5] in 1949 treat the numbers of individuals of each class(**S**, **I**, and **R**) as stochastic variables as will be seen in **Section 2**. He developed a time evolution

equation taking into account probability distributions on the number of individuals of each class.

(2)Stochastic approach.

It will be assumed from the beginning that the disease propagation is, directly or indirectly, due to interactions between individuals of the population. This implies that stochastic effects would be *responsible* for the epidemic spreading. In this way the number n of individuals can be taken as the primary stochastic variables to build a theoretical approach in order to explain the plague evolution. Thus, let us assume that the population has a very large number N of individuals formed by three different classes of persons: n_1(**susceptible**), n_2(**infected**) and n_3 (**recovered**) and that $n_1+n_2+n_3 = N$ = constant. These hypothesis are valid when the epidemic duration is relatively short so that one can neglect the population modification due to births or to deaths with diseases different from the epidemic.

(2.a) Master equation.

In a recent paper[6], analyzing the infected number n(t) by the covid19 in Italy, using Langevin equation, we verified that stochastic effects plays a significant role in the plague propagation. Stochastic effects now will be investigated using the **Master Equation**.[7,8]

So, let us indicate by P(n,t) the probability to find at a time t, persons of kind n. Due to recovering and infections, the time evolution of P(n,t) will be described by the Master Equation:[7,8](see Appendix A)

$$dP(n,t)/dt = \Sigma_r \Sigma_{n'} \{W_r(n|n')P(n',t) - W_r(n'|n)P(n,t)\} \qquad (2.1),$$

where $W_r(n|n')$ are the transition rates between the different classes, that is, between the state n and the states n´, where n´= n_1, n_2 and n_3.

For infecting transition **S** → **I** ($n_1 \rightarrow n_2$), when n_1 decreases by one unit and n_2 remains invariant, we have

$$W_{inf} = W_{12} = -b\, N(n_1/N)(n_2/N), \qquad (2.2)$$

where b is the infection rate constant.

For recovering transition **I**→**R** ($n_2 \rightarrow n_3$), when the infected number n_2 decreases by one unit and the recovered n_3 increases by one unit we have

$$W_{rec} = W_{23} = c\, N(n_2/N) \qquad (2.3),$$

where c is the recovery rate constant.

As $n_1 + n_2 + n_3 = N$,

$$dn_1/dt + dn_2/dt + dn_3/dt = 0,$$

and consequently,

$$\begin{aligned} (\mathbf{S\rightarrow I}) \qquad & dn_1/dt = -b\, N(n_1/N)(n_2/N), \\ & dn_2/dt = b\, N(n_1/N)(n_2/N) - c\, N(n_2/N) \quad (2.4) \quad \text{and} \\ (\mathbf{I\rightarrow R}) \ldots\ldots & dn_3/dt = c\, N(n_2/N). \end{aligned}$$

(2.b) Fokker-Planck equation.[8,9]

Taking into account that $N >> n_i$ transitions rates in the Master Equation (2.1), that were estimated in terms of the differences $n' - n = \pm 1$ or 0, can be now calculated in terms of very small differences $x - x' \approx 1/N = \varepsilon << 1$. This allows us to expand the quantities on the right-hand side of Eq.(2.1) around x getting,

$$dP(x,t)/dt = N\Sigma_r\Sigma_{x'}\{\omega_r(x|x')P(x',t) - \omega_r(x'|x)P(x,t)\} \qquad (2.5),$$

where $x \equiv (x_1,x_2,x_3)$, $\omega_r = W_r/N$; $\omega_1 = -bx_1x_2$, $\omega_2 = bx_1x_2 - cx_2$ and $\omega_3 = cx_2$.

Note that in this context Eqs.(2.4) becomes written as

$$\begin{aligned} & dx_1/dt + dx_2/dt + dx_3/dt = 0, \\ (\mathbf{S\rightarrow I}) \quad & dx_1/dt = -bx_1x_2, \\ & dx_2/dt = bNx_1x_2 - cx_2 \quad \text{and} \qquad (2.6), \\ (\mathbf{I\rightarrow R}) \quad & dx_3/dt = cNx_2\ . \end{aligned}$$

Noting that the transition probabilities ω_r involve only two independent variables at each time, like (x_1, x_2) and (x_3,x_2), performing[4] the expansion of the second member of Eq.(2.5) up to second order in ε we obtain (see Appendix B) the **Fokker-Planck**[7] equation:

$$\partial P(x,t)/\partial t = -\Sigma_i\, \partial(f_iP)/\partial x_i + (\varepsilon/2)\, \Sigma_{ij}\, \partial^2(f_iP)/\partial x_i\partial x_j \qquad (2.7).$$

where $f_i = \Sigma_r\, \omega_r$ (r =1,2,3), with $\omega_1 = -bx_1x_2$, $\omega_2 = bx_1x_2 - cx_2$ and $\omega_3 = c\, x_2$.

(2.c)Evolution of the averages.

Solving Fokker-Planck equation (2.5), taking into account that in the transitions $x \rightarrow x'$, $x - x' \sim \varepsilon$, we obtain P(x,t) that is centered around $x \equiv (x_1,x_2)$. With P(x,t) we calculate the averages:

$$< x_i(t) > = \iint x_i\, P(x)\, dx \qquad (2.8).$$

Now, multiplying both sides of the Fokker-Planck Eq.(2.7) by x_i and integrating in x results[4]:

$$d< x_i >/dt = < f_i > = < \omega_i > \qquad (2.9),$$

where $\omega_1 = -bx_1x_2$, $\omega_2 = bNx_1x_2 - cx_2$, and $\omega_3 = cx_2$.

To get Eq.(2.9) we have performed appropriate integration by parts and considered that P(x,t) vanishes quickly as the limits of integral is approached.[4]

For $\varepsilon \rightarrow 0$ the probability distribution P(x,t) becomes sharped Gaussians centered in x_i.[4] In this way, we have $<x_i x_j> \approx < x_i >< x_j >$. Defining, $x(t) = < x_1(t) >$, $y(t) = < x_2(t) >$, $z(t) = < x_3(t) >$ and remembering that and $x(t) + y(t) + z(t) = N$ = constant, Eqs.(2.6) becomes written as:[4]

$$\begin{aligned} dx/dt + dy/dt + dz/dt &= 0, \\ dx/dt &= -bNxy, \\ dy/dt &= bNxy - cy \quad \text{and} \quad (2.10), \\ dz/dt &= cNy \end{aligned}$$

(3)Epidemic outbreak.

The outbreak of an epidemic phenomenon is characterized being a critical event. If the number of infective individuals is small there is no spread of the disease. But if this number increases it will reach a critical value above which the epidemic spreads, the increase of the infectious persons being exponential in time. This fundamental idea was used by Ross[2] in his studies on the transmission of malaria and was introduced by Kermack and McKendrick[2] in a clear form as the *threshold theorem*. According to Mario and Tania[4] analysis, in the epidemic outbreak regime the evolution equations (2.10) are satisfied.

(4) Examples of epidemic models.

Remembering:

S(t) = number of individuals not yet infected with disease at time t, or those susceptible to the disease of the population.

I(t) = number of individuals of the population infected with the disease and that are capable of spreading the disease to those in the **S** category.

R(t) = number of individuals of the population that have been infected and them removed from **S**(t), either due to immunization or due to death. They are not able to be infected again or to transmit the infection to others.

Many models have been proposed[2,4] to describe epidemics. They are known, for instance, by the names **SIR,SIS, SEIR** and **SIRS**.

(4.a) SIR model.

In this model, indicated by **S** →**I** →**R** the time epidemic evolution is described by the Eqs.(2.4), remembering that x+y+z = 1:

$$\begin{aligned} &(\mathbf{S}\rightarrow\mathbf{I}) \quad dx/dt = -bNxy \\ &\qquad\qquad dy/dt = bNxy - cy \qquad \text{and} \qquad (4.1), \\ &(\mathbf{R}\rightarrow\mathbf{I}) \quad dz/dt = cNy \end{aligned}$$

where x(t) = **S**(t), y(t)= **I**(t) and z(t) = **R**(t). In the reaction **S** →**I** the number **S** decreases by one unit and **I** remains invariant. In the reaction **R** → **I,** the number **I** decreases by one unit and **R** increases by one unit, according to Eqs(2.4).

(4.b) SIS model.

In this model the flux is indicated by **S** → **I** and **I** → **S** and the time evolution disease obeys the differential equations, where x + y = 1:

$$\begin{aligned} &(\mathbf{S}\rightarrow\mathbf{I}) \quad dx/dt = -bxy + cy \\ &(\mathbf{I}\rightarrow\mathbf{S}) \quad dy/dt = bxy - cy \qquad (4.2). \end{aligned}$$

(4.c) SEIR model.

Some **S** individuals that have been infected takes a certain time to be infective. These individuals **E** ("exposed") are not able to infect others. The flow **S** → **E**→ **I**→ **R** is governed by the equations

$$\begin{aligned} &(\mathbf{S}\rightarrow\mathbf{I}) \quad dx/dt = -bxy \\ &\qquad\qquad de/dt = bxy - ke. \qquad (4.3) \\ &\qquad\qquad dy/dt = ke - cy \\ &(\mathbf{R}\rightarrow\mathbf{I}) \quad dz/dt = cy \end{aligned}$$

where the constant k is responsible for the process **E** →**I** and Eqs.(4.3) are not all independent because x + e + y+ z = 1.

(4.d) SIRS model.

The flux is given by **S**→ **I** →**R** →**S** and the rate equations are

$$dx/dt = -bxy + az$$

$$dy/dt = bxy - cy \qquad (4.4),$$
$$dz/dt = cy - az$$

that are not all independent because x + y + z = 1.

(4.e) Estimations of Spreading evolution and Conclusions.

Spreading evolution is obtained determining the functions x(t),y(t) and z(t), integrating numerically the linear differential functions defined by the Eqs.(2.10) or from simulations of the **Master Equation** (2.1).

This was done, for instance, by Tânia and Mario.[4] They shown that the stochastic predictions obtained with the numerical integrations of the linear differential equations involving x(t)= **S(t),** y(t) **= I(t)** and z(t) = **R(t).** are very similar to that obtained with the Master Equation simulations.

From this one can conclude that a good estimation of the epidemic spreading can be performed with the employment of ordinary differential-equations of the first order in time, according to the **deterministic models.**

(5)Deterministic models

Till 1949 many epidemic models have been proposed employing ordinary differential-equations of the first order in time, to describe the relationship between **Susceptible(S)**, **Infected(I)** and **Recovered(R)** individuals in a population.[2]These theoretical studies became known as the **deterministic approach** of the epidemic spreading[3] However, this approach did not describe, in a explicit manner, the random fluctuations occurring in a real epidemic spreading. In order to take into account random fluctuations, Bartlett[2,4] in 1949 treat the numbers of individuals of each class(**S**, **I**, and **R**) as stochastic variables as seen in Sections 2-4.

According to the **Section (5.c)** we concluded that the deterministic approach gives a good description of the epidemic spreading. Many deterministic models SIR, SIS,SISR,SIRI,SIRS and SEIR are analyzed, for example, in reference[3].We show here only the **SIR** and **SIS** models.

(5.a)SIR model.

The disease flow of this model is indicated by **S** → **I** → **R**. In this model, dS/dt, dI/dt and dR/dt are given by

$$dS/dt = -\beta SI/N$$

$$dI/dt = \beta SI/N - \gamma I. \qquad (5.1)$$

$$dR/dt = \gamma I.$$

In the reactions **S→I** the number **S** decreases by one unit and **I** remains invariant. In the reactions **R → I**, the number **I** decreases by one unit and **R** increases by one unit. The population number N =constant = **S+I+R**, that is ,the rate of infections and recovery is much faster than time scale of births and deaths. The parameters β and γ are constant transmission probability factors per unit of time.[2]

In this case **S**(t)+**I**(t)+**R**(t) = N = constant. Solving numerically Eqs.(5.1) we get the time evolution of S(t), I(t) and R(t). In **Figure 2**[2] these functions are shown for the initial values S(0) = 997, I(0) = 3, R(0) = 0, β = 0.4 and γ = 0.04.

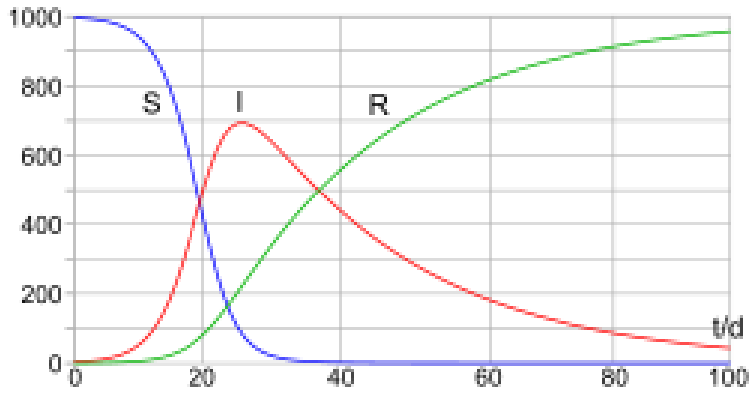

Figure 2. Functions S(t), I(t) and R(t) for the **SIR** model. Time t is measured in days.

This **SIR** case, shown in Fig.2, should have an ideal epidemic evolution: the infected individuals decreases quickly tending to zero. The disease does not become endemic.

(5.b)SIS model.

For this model the flow is represented by **S →I →S** and the time disease evolution are governed by the equations

$$dS/dt = -\beta SI/N + \gamma I$$

$$dI/dt = \beta SI/N - \gamma I. \qquad (5.2).$$

Also in this case **S +I + R** = constant= N.

(5.c) SEIR model.

Some **S** individuals that have been infected takes a certain time to be infective. These individuals **E** ("exposed") are not able to infect others. The flow **S** → **E**→ **I**→ **R** is governed by the equations

$$\begin{aligned} dS/dt &= -\beta SI/N \\ dE/dt &= \beta SI/N - kE. \\ dI/dt &= kE - \gamma I. \\ dR/dt &= \gamma I \end{aligned} \qquad (5.3),$$

where the constant k is responsible for the process **E** →**I** and the Eqs.(5.3) are not all independent because **S + E + I+ R** = N.

(6.d) SIRS model.

The flux is given by **S**→ **I** →**R** →**S** and the rate equations are

$$\begin{aligned} dS/dt &= -\beta SI/N + \alpha R \\ dI/dt &= \beta SI/N - \gamma I \\ dR/dt &= \gamma I - \alpha R \end{aligned} \qquad (5.4),$$

that are not all independent because **S + I + R =** N.

Figures showing S(t), I(t) and R(t) as functions of t can be seen, for instance, in references [3] and [4].

APPENDIX A. Master Equation.[7]

Master equation is applied, for example, to describe chemical reactions when molecules of different kinds are transformed due to interactions between them. Let us assume that $P_j(t)$ is the probability to find a molecule j at time t. Usually, the transition rates describing these chemical reactions are represented by a matrix **A**. In this way, the transition rate (transition per unit of time) dP_k/dt of a molecule k to be appear in the reactions with molecules j would be given by the Master Equation(ME)

$$dP_k/dt = \sum_j A_{kj} P_j \qquad (A.1),$$

where P_k is the probability to find a molecule of j kind and the matrix **A** is filled with parameters that describe the j→k transitions. The ME can be simplified so that the terms with j = k do not appear in the summation:

$$dP_k/dt = \sum_{j\neq k} \{A_{kj} P_j - A_{jk} P_k\} \qquad (A.2)$$

APPENDIX B. Fokker-Planck Equation[8,9]

The Fokker-Planck (FP) equation is a partial differential equation that describes the time evolution of the probability density functions that can obtained by adequate transformations of the Master Equation. For instance, by the Kramers-Moyal expansion.

In our case the Fokker-Planck equation will be obtained from the Master Equation (2.5)[4] by a Taylor series[10] expansion in a second order of $\varepsilon = x' - x$. Since $\omega_r = f(x_i, x_j)$, Eq.(2.5) is written as,

$$dP(x)/dt = N\Sigma_r\Sigma_{x'}\{\omega_r(x_i,x_j|x_i', x_j')P(x') - \omega_r(x_i', x_j'|x_i,x_2)P(x)\} \qquad (B.1)$$

First, let us remember that a function of two variables f(x,y) expanded around two points $x + \varepsilon_x$ a and $y + \varepsilon_y$ is given by,[10]

$$f(x +\varepsilon_x, y +\varepsilon_y) \approx f(x,y) +f_x(x,y)\varepsilon_x +f_y(a,b)\varepsilon_y +f_{xx}(a,b)\varepsilon_x^2/2 + f_{yy}(a,b)\varepsilon_y^2/2 +f_{xy}(a,b)\varepsilon_x\varepsilon_y \qquad (B.2).$$

Defining $f(x_i',x_j') = \omega_r(x_i',x_j'|x_i, ,x_j)P(x)$ and $f(x_i, x_j)= \omega_r(x_i, x_j|x_i', x_j')P(x')$, putting $x_i' = x_i+ \varepsilon_i$, $x_j' = x_j + \varepsilon_j$ and $\varepsilon_i = \varepsilon_j = \varepsilon = 1/N$, Eq.(B.1) becomes given by

$$\partial P(x,t)/\partial t = -\Sigma_i \partial(f_iP)/\partial x_i + (\varepsilon/2)\, \Sigma_{ijr}\, \partial^2(f_iP)/\partial x_i\partial x_j \qquad (B.3).$$

Acknowledgements.

The author thanks M.J.de Oliveira for the unpublished manuscripts.

REFERENCES

[1]Pandemic - https://en.wikipedia.org/wiki/Pandemic
[2] Mathematical modeling of infectious disease - https://wikipedia.org/wiki/Mathematical
[3] "Epidemic spreading". T.Tomé and M. J. de Oliveira. Rev.Bras.Ens.de Física (to be published).
[4] "Stochastic approach to the epidemic spreading" T.Tomé and M. J. de Oliveira. Braz. J. of Physics (to be published).
[5]M.S.Bartlett. *Stochastic Processes,* University of North Caroline(1947). J.R.Stat.Soc.B11,211(1949).
[6]M.Cattani, M.C.Salvadori and F.S.Teixeira. http://publica-sbi.if.usp.br/PDFs/pd1719.pdf
[7] Master Equation. https://en.wikipedia.org/wiki/Master_equation
[8]Fokker-Planck equation. https://en.wikipedia.org/wiki/Fokker-Planck_equation https://arxiv.org/pdf/cond-mat/0701242.pdf
[9]"Dinâmica Estocástica e Irreversibilidade". T.Tomé and M. J. de Oliveira. Edusp(2014).
[10]Taylor Series Expansion https://math.libretexts.org/Bookshelves/Calculus/Supplemental_Modules_(Calculus)/Multivariable_Calculus/3%3A_Topics_in_Partial_Derivatives/Taylor__Polynomials_of_Functions_of_Two_Variables