ASIAN JOURNAL OF STATISTICS AND APPLICATIONS

VOL NO. 1, ISSUE 2, PAGE NO. 71-78 YEAR 2024

An Application of Stochastic Modelling to COVID-19 Infections data of India

SEAL, Babulal[a], BASU, Sudipta[b]

[a] *Department of Mathematics and Statistics, Aliah University, New Town, Kolkata*
[b] *Department of Mathematics and Statistics, Aliah University, New Town, Kolkata*

ARTICLE HISTORY
Online Published December 27, 2024

ABSTRACT
In this article it is tried to construct a stochastic model which can be used to describe biological growth phenomena that roughly shows an exponential pattern but there is a resistance force(could be manual) acting on it.It is mainly constructed to explain growth dynamics of infected patients by COVID-19 in those countries where an early lockdown was imposed.Here it is attempted to find the expression of variable of interest at time t and also the MLEs of model parameters are worked out.This model is applied to a real life data of infected patients by COVID-19 in India, specially for the period in which there was a strict lockdown around the country. Then a comparative study is made with exponential model that could be a possible option to describe such situation. This model could be used to the data of any country, where an early lockdown decision was made as a precautionary measure to deal with this situation.This article aims to see whether there is any impact of the resistant force or not.

KEYWORDS
COVID-19, Ito's Lemma, Growth deceleration factor, Exponential model.

Introduction

There are some variants of Coronavirus that exists and scientists classified them into four subgroups-1)229E(alpha), 2)NL63(alpha), 3)OC43(beta), 4)HKU1(beta) and there are three rare types also and they are- 1)MERS-CoV, 2)SARS-CoV, 3)SARS-CoV-2 ([11]). The third one under the rare category which is also known as COVID-19. It is mainly responsible for the pandemic situation that occured all around the world. It is actually an infectious disease. It is observed that most of the infected persons by this virus, experienced a moderate respiratory problems and many of them recovered under normal treatments, but for older persons who were suffering from CVD(cardio vascular disease), severe respiratory problems, diabetes, cancer etc. had faced serious illness after this virus infection([10]). As this is an infectious disease, it is natural to be interested about the manner it spreads. This virus(COVID-19) actually spreads itself via droplets of saliva or when a person infected by this virus sneezes or coughs. There is every possibility that surrounding persons may become infected by it(WHO report,2020). One can classify this virus as a respiratory virus and it is important to have an idea about the possible ways of transmission of these respiratory virus. Sources

CONTACT Author[a]. Email:babulal.seal@gmail.com

indicate that there are mainly three ways of transmission of such respiratory virus, firstly it could be transmitted via direct contact with infected person, secondly through droplet transmission and thirdly, through airborne transmission of smaller droplets and particles that stays in air for fair amount of time and can travel a significant amount of distance[13], not only this, some Italian scientists collected outdoor air pollution samples and after studying it, they found gene highly specific to COVID-19 in multiple samples(Coronavirus detected on particles of air pollution, The Guardian) i.e, one may think it as airborne, though there is not much evidence of this fact. But one can consider it as partially airbone because of the fact that air current can help it to travel from one place to another and because of this reason, it is quite possible that some of the countries like South Korea, Sweden etc, haven't imposed any strict lockdown or curfew. So far from the experience of patients who had an infection of this COVID-19 virus, it is described that there are some specific symptoms through which a person may be infected by this virus. The symptoms are fever(high temperature), cough, breathing problem, headache, sore throat, loss of smell or taste([10]). Observing the situations all around the world, one can understand the severity of infection.

Like most of the biological growth phenomena, growth dynamics of number of infected patients shows exponential pattern until some resistance force generated by nature comes into play. To prevant the infection, most of the countries including India have taken few precautionary steps like social distancing, home quarantine, curfew etc. The motivation behind construction of such a model(eq. 1) comes from the fact that if a considerable amount of people are infected by this virus, then they can infect a fairly large amount of people even if there are precautionary measures ([13]). On the other hand few countries had their belief in herd immunity theory which states that it is a form of indirect protection from infection which occurs when a big part of the population has become immune to the infection([9]). Keeping this in mind and observing the pattern of RGR of the data of India, it is reasonable to consider the equation,

$$\frac{dx(t)}{dt} = ax(t) - b(1/x(t))^p \tag{1}$$

where, a is growth rate factor and b is growth resistance factor. The motivation behind the construction of this model is that the countries like India where an early lockdown decision was made to control the number of infected patients and it is purely a manual attempt. It is observed that this virus is airbone(WHO,2020). So, if the number of infected patients increase abruptly, then the situation of community spread can occur, i.e, the value of x(t) will be higher and higher and it is easy to understand that the growth resistant factor will become inactive i.e, there will be only growth rate factor which will be active untill herd immunity is achieved.

This article is constructed in the following sections. In section 2, a stochastic model is constructed by taking the variability of the growth rate factor a. Also the MLEs of the model parameters are worked out. In section 3, a generalization of the model of section 2 is constructed. The MLEs of the model parameters are worked out. In section 4, analysis of the COVID-19 infection data of India is presented. In section 5, Some Conclusion are made.

1. Stochastic Extension of the model

It is already stated that except for few countries, who believed in herd immunity theory, most of the countries imposed precautionary measures to slow the rate of increase of number of infected patients. Now, it is obvious that environmental fluctuations and atmospheric changes may have a significant impact on growth rate factor also. During the process there are few oscillations occured and to deal with it we introduce multiplicative noise term. On the other hand this noise term can also be introduced as many authors([2],[8]) considered this variation with time of model parameter by a rule, which is mentioned as,

$$\theta(t) = a + \sigma.\eta(t)$$

which takes us to the stochastic version of the model described by equation as,

$$dx(t) = (ax(t) - b(1/x(t))^p)dt + \sigma.x(t)dw(t) \tag{2}$$

where w(t) is the standard Gaussian process and the differential is meant in Ito's sense([1]). Now, as it appears that it is a SDE in Ito's form, one can solve it using Ito's lemma. Using the transformation,

$$y(t) = F(x,t) = x^{p+1}$$

$$F'(x,t) = (p+1)x^p$$

$$F''(x,t) = p(p+1)x^{p-1}$$

Applying Ito's lemma we get,

$$f = (p+1)(a + \frac{\sigma^2 p}{2})y - b(p+1)$$

$$g = \sigma.(p+1)y$$

This forms reduced SDE as,

$$dy(t) = [(p+1)(a + \frac{\sigma^2 p}{2})y(t) - b(p+1)]dt + \sigma.(p+1)y(t)dw(t)$$

Now above is a SDE in standard form and there are standard method available(Linda.J.S.Allen,2010) to solve such SDE.Using the standard method the solution becomes,

$$y(t)=exp([\frac{a(p+1)+\sigma^2 p(p+1)}{2}]t+\sigma.(p+1)w(t))[y(0)-b(p+1)\int_0^t exp([-b(p+1)-a(p+1)+\frac{\sigma^2 p(p+1)}{2}]s-\sigma.(p+1)w(s))ds]$$

Transforming it to the original variable x(t),

$$x(t)=(exp([\frac{a(p+1)+\sigma^2 p(p+1)}{2}]t+\sigma.(p+1)w(t))[(x(0))^{p+1}-b(p+1)\int_0^t exp([-b(p+1)-a(p+1)+\frac{\sigma^2 p(p+1)}{2}]s-\sigma.(p+1)w(s))ds])^{\frac{1}{p+1}} \tag{3}$$

MLEs of the model parameters

Now to obtain the MLE of a and b on the basis of continuous records available one may assume that the diffusion coefficient terms are known. So this part does not involve any unknown parameters.Using Girsanov theorem([4],[5]) the infill log-likelyhood becomes,

$$l(a,b) = A - \frac{1}{2}B$$

Where,

$$A = \int_0^T \frac{ax(t) - b\frac{1}{x(t)^p}}{\sigma^2(x(t))^2} dx(t)$$

$$B = \int_0^T \frac{(ax(t) - b\frac{1}{x(t)^p})^2}{\sigma^2(x(t))^2} dt$$

The MLEs of **a** and **b** are given as,

$$\hat{a} = \frac{lnx(T)}{T} + \frac{1}{T}I_1 \frac{\frac{lnx(t)}{T}I_1 - I_2}{I_3 - \frac{1}{T}I_4}$$

where,

$$I_1 = \int_0^T \frac{1}{(x(t))^{p+1}} dt,$$

$$I_2 = \int_0^T \frac{1}{(x(t))^{p+2}} dx(t),$$

$$I_3 = \int_0^T \frac{1}{(x(t))^{2p+2}} dt,$$

$$I_4 = (\int_0^T \frac{1}{(x(t))^{p+1}} dt)^2$$

and

$$\hat{b} = \frac{\frac{lnx(T)}{T}J_2 - J_3}{J_4 - \frac{1}{T}J_5}$$

where

$$J_1 = \int_0^T \frac{1}{(x(t))^{p+1}} dt,$$

$$J_2 = \int_0^T \frac{1}{(x(t))^{p+2}} dx(t),$$

$$J_3 = \int_0^T \frac{1}{(x(t))^{2p+2}} dt,$$

$$J_4 = (\int_0^T \frac{1}{(x(t))^{p+1}} dt)^2$$

2. Time dependent model

As it mentioned earlier that this articles mainly focuses on those countries where an early lockdown was imposed like India. The variability of growth rate factor with time is already done. Now in this section the variability of both the factors are considered. Lockdown is purely a manual step to control the number of infected cases but this manual step also changed it's intensity from time to time according to the need of the current situation. So, we incorporate the variability of both the factors with time by the following rule which was introduced by some authors([2],[8]) as,

$$\theta(t) = a + \sigma_1 \eta_1(t)$$

$$\mu(t) = b + \sigma_1 \eta_1(t)$$

It enables us to consider generalized stochastic version of equation 1 as,

$$dx(t)[ax(t) - b(\frac{1}{x(t)})^p]dt + [\sigma_1 x(t) - \sigma_2(\frac{1}{x(t)})^p]dw(t) \tag{4}$$

It is again a SDE in Ito's from and therefore, it needs to be solved by Ito's lemma. So, applying the following transformation,

$$F(x,t) = x^{p+1}$$
$$F'(x,t) = (p+1)x^p$$
$$F''(x,t) = p(p+1)x^{p-1}$$

and applying Ito's lemma we get,

$$f = (p+1)[ay - b] + \frac{p(p+1)}{2}[\sigma_1^2 y - 2\sigma_1\sigma_2 + \frac{\sigma_2^2}{y}]$$

$$g = \sigma_1 y - \sigma_2$$

If one looks at the data set, the values are large in nature and because of this reason, the reduced SDE can be well approximated as,

$$dy(t) \approx [(a(p+1) + \frac{\sigma_1^2 p(p+1)}{2})y(t) - (b(p+1) + \sigma_1\sigma_2 p(p+1))]dt + (\sigma_1 y(t) - \sigma_2)dw(t)$$

Just to simplify above SDE, another transformation is applied as

$$F(y,t) = y - \frac{\sigma_2}{\sigma_1}$$

$$F'(y,t) = 1$$
$$F''(y,t) = 0$$

Therefore reduced SDE becomes,

$$dz(t) = [(a(p+1) + \frac{\sigma_1^2 p(p+1)}{2})z(t) - (b(p+1) + \sigma_1\sigma_2 p(p+1)) + \frac{\sigma_2}{\sigma_1}(a(p+1) + \frac{\sigma_1^2 p(p+1)}{2})]dt + \sigma_1 z(t)dw(t)$$

Above is a SDE in standard form and there are standard methods available([3],[6]) to yield the solution. Using standard method the solution of the SDE becomes,

$$z(t) = exp([\frac{a(p+1) + \sigma_1^2 p(p+1)}{2}]t + \sigma_1(p+1)w(t))[z(0) + \frac{\sigma_2}{\sigma_1}(a(p+1) + \frac{\sigma_1^2 p(p+1)}{2})\int_0^t exp([-\frac{\sigma_1^2 p(p+1)}{2} - \frac{a(p+1)}{2}]s - \sigma_1 w(s))ds]$$

Now, transforming it to original variable we get,

$$x(t) = [\frac{\sigma_2}{\sigma_1} + exp([\frac{a(p+1) + \sigma_1^2 p(p+1)}{2}]t + \sigma_1(p+1)w(t))[((x(0))^{p+1} - \frac{\sigma_2}{\sigma_1}) + \frac{\sigma_2}{\sigma_1}(a(p+1) + \frac{\sigma_1^2 p(p+1)}{2})\int_0^t exp([-\frac{\sigma_1^2 p(p+1)}{2} - \frac{a(p+1)}{2}]s - \sigma_1 w(s))ds]]^{\frac{1}{p+1}}$$

MLEs of model parameters

Here also to obtain the MLE of a and b on the basis of continuous records available one may assume that the diffusion coefficient terms are known. So this part does not involve any unknown parameters.Using Girsanov theorem([4],[5]) the infill log-likelyhood becomes,

$$l(a,b) = A - \frac{1}{2}B$$

where,

$$A = \int_0^T \frac{ax(t) - b\frac{1}{x(t)^p}}{f(x(t))}dx(t)$$

and

$$B = \int_0^T \frac{(ax(t) - b\frac{1}{x(t)^p})^2}{f(x(t))}dt$$

The MLEs of a and b are

$$\hat{a} = \frac{A_2A_5 - A_3A_4}{A_2^2 - A_1A_4}$$

where,

$$A_1 = \int_0^T \frac{(x(t))^2}{f(x(t))}dt,$$

$$A_2 = \int_0^T \frac{(x(t))^{1-p}}{f(x(t))}dt,$$

$$A_3 = \int_0^T \frac{x(t)}{f(x(t))}dx(t),$$

$$A_4 = \int_0^T \frac{(x(t))^{-2p}}{f(x(t))}dt,$$

$$A_5 = \int_0^T \frac{(x(t))^{-p}}{f(x(t))}dx(t)$$

and

$$\hat{b} = \frac{B_1B_5 - B_2B_3}{B_2^2 - B_1B_4}$$

with

$$B_1 = \int_0^T \frac{(x(t))^2}{f(x(t))}dt,$$

$$B_2 = \int_0^T \frac{(x(t))^{1-p}}{f(x(t))}dt,$$

$$B_3 = \int_0^T \frac{x(t)}{f(x(t))}dx(t),$$

$$B_4 = \int_0^T \frac{(x(t))^{-2p}}{f(x(t))}dt,$$

$$B_5 = \int_0^T \frac{(x(t))^{-p}}{f(x(t))}dx(t)$$

3. Analysis of real life data

Here, a real life data set of no. of infected persons in India by COVID-19 at different time points are taken(data source:worldometer).In this data set of infected persons at every 7th day are given.Data points are available from April 03,2020 to July 03,2020. It is known from social media sources, news channels and news papers that most of the countries have their belief in lockdown theory. Here, it is tried to fit this proposed model with p equals 1(growth rate factor varies with time) to the data set on the basis of the RGR graph and the following results are obtained,

$$\hat{a} \approx 0.538,$$

$$std.err \approx 0.09,$$

$$\hat{b} \approx 1,$$

$$std.err \approx 0.001,$$

and

$$AIC \approx 302.38$$

One can see that a comparative study with exponential model is to be conducted here because the growth pattern suggests so. If one observes the picture of the RGR it will be understood that exponential model will not be an appropriate model and this fact is evidenced graphically.

On the other hand Exponential model does not posses any resistance force. Still one may look at the results of Exponential model given by

$$\hat{a} \approx 0.54,$$
$$std.err \approx 0.09,$$

and

$$AIC \approx 302.38$$

From this we see that there is no much improvement in the introduced model than the exponential model. It is mainly the first model(Eq. 2) that is tried to fit to the data and compared with the Exponential one. The generalized version is worked out here and to see whether there are resistant factors of other forms. If it is so and varies with time the generalized version may play an important role to describe that particular process.

4. Conclusion

As it is already dicussed about the ideology behind the construction of the proposed model, one can understand the applicability of this model to the growth dynamics of different Biological systems where the graph shows an exponential pattern with a resistance force acting on the process. This proposed model is applicable not only for Biological system but to the other areas also. One can asses the performance and compare it with other existing models whenever such datasets are available. It is way of generalizing the model.

References

[1] R.S. Liptser, and A.N. Shiryayev., *Statistics of random process . I.General Theory*,Newyork :Springer-verlag, 1977.
[2] L. Ferrante, S. Bompadre, L. Possati, and L. Leone., *Parameter Estimation in a Gompertzian Stochastic Model for Tumor Growth*,Journal Biometrics,Vol-56, 2000.
[3] B. Oksendal., *Stochastic Differential Equations*,Springer, 2006.
[4] Jaya.P.N. Bishwal., *Parameter Estimation in Stochastic Differential Equation*,Springer, 2008.
[5] P.C.B. Phillips, Jun. Yu., *Maximum likelihood and Gaussian estimation of continuous time models in finance*, Handbook of financial time series, 2009.
[6] Linda.J.S. Allen., *An Introduction to Stochastic Processes with Application to Biology*, CRC press, 2010.
[7] Michael.J. Panik., *Growth Curve Modeling;with applications*,Wiley, 2013.
[8] S. Basu, B. Seal, *A Modified Gompertzian Stochastic Model With Application*, Journal of Agricultural and Statistical sciences,Vol-15, 2019
[9] Wikipedia(updated,2020), Corona virus.
[10] WHO(2020), Report.
[11] Webmd(updated,2020), Corona virus;disease,symptoms.
[12] The Guardian(2020), Corona virus detected on particles of airpollution.
[13] Editorial Report(2020) COVID-19 transmission-up in the air, The Lancet Respiratory Medicine. integral function,

Appendix

```
library(Sim.DiffProc)
x=c(data)
length(x)
X=as.ts(x)
```

```
t=1:14
plot(t,x,xlab="time",ylab="no. of infected persons",type="l")
r=numeric(13)
for(i in 1:13)
r[i]¡-(x[i+1]-x[i])/x[i]

r
plot(r,type="l")
fx = drift term
gx = diffusion term

fitmod1 = fitsde(X,fx,gx,start = list(theta1=1,theta2=1, theta3=1),optim.method =
"L-BFGS-B")

fitmod1
summary(fitmod1)
```