MANAS Journal of Engineering
ISSN 1694-7398 | e-ISSN 1694-7398

Volume 9, Issue 2, (2021) Pages 192-197
https://doi.org/10.51354/mjen. 929656

Fitting the Itô Stochastic differential equation to the COVID-19 data in Turkey

Sevda Özdemir Çalıkuşu[1*], Fevzi Erdoğan[2]

[1] Yuzuncu Yıl Unıversıty, Özalp Vocational School, Turkey, sevdaozdemir@yyu.edu.tr, ORCID: 0000-0002-0238-2677

[2] Van Yuzuncu Yil University, Faculty of Economics and Administrative Sciences, Turkey, ORCID: 0000-0003-3745-0198

ABSTRACT

In this study, COVID-19 data in Turkey is investigated by Stochastic Differential Equation Modeling (SDEM). Firstly, parameters of SDE which occur in mentioned epidemic problem are estimated by using the maximum likelihood procedure. Then, we have obtained reasonable Stochastic Differential Equation (SDE) based on the given COVID-19 data. Moreover, by applying Euler-Maruyama Approximation Method trajectories of SDE are achieved. The performances of trajectories are established by Chi-Square criteria. The results are acquired by using statistical software R-Studio.These results are also corroborated by graphical representation.

ARTICLE INFO

Research article

Received: 30.04.2021
Accepted: 20.10.2021

Keywords:
Itô stochastic differential equation, Euler-Maruyama approximation method, Maximum likelihood estimation method, COVID-19 data

[*]Corresponding author

1 Introduction

Coronavirus disease (COVID-19) is an infectious disease caused by a newly discovered coronavirus. Most people infected with the COVID-19 virus will experience mild to moderate respiratory illness and recover without requiring special treatment. Older people, and those with underlying medical problems like cardiovascular disease, diabetes, chronic respiratory disease, and cancer are more likely to develop serious illness [1]. The first case of COVID-19 in Turkey was reported on 11 March 2020, and the number of reported cases has increased day by day.

Stochastic differential equations model (SDEM) stochastic evolution as time evolves. These models have a variety of applications in many disciplines and emerge naturally in the study of many phenomena. Examples of these applications are physics, astronomy, mechanics, economics, mathematical finance, geology, genetic analysis, ecology, cognitive psychology, neurology, biology, biomedical sciences, epidemiology, political analysis and social processes, and many other fields of science and engineering [2]. And it is known that SDEMs are more realistic mathematical model than normal differential equation models of the situation [3]. Because stochastic differential equations, unlike normal differential equations, contain random effects called "white noise".

In [4], Kostrista E. and Cibuku D. are looking for answers to some questions which are related with stochastic differential equations. These questions are 'In what situations does the SDE arise?', 'What are its essential features?', 'What are the applications and the connections to the other fields?' and 'How does SDEs model the physical situation and white noise process which is the generalized mean-square derivative of the Wiener process or Brownian motion?'. In [5], Ince N. and Shamilov A. develop a new method to obtain approximate probability density function (pdf) of random variable of solution of stochastic differential equations (SDEs) by using generalized entropy optimization methods (GEOM). By starting given statistical data and Euler–Maruyama (EM) method approximating SDE are constructed several trajectories of SDEs. The constructed trajectories allow to obtain random variable according to the fixed time. In [6], Mahrouf M. et al. extend the well-known SIR compartmental model to deterministic and stochastic time-delayed models in order to predict the epidemiological trend of COVID-19 in Morocco and to assess the potential role of multiple preventive measures and strategies imposed by Moroccan authorities. In [7], Bak J. Et al. propose a method to test for lack-of-fit of an estimated stochastic differential equation. The method is based on Monte Carlo simulation of trajectories between neighbour observations and, thus, it does not rely on the availability of explicit expressions of the conditional densities. Consequently, both non-linear models and models with state-dependent drift and diffusion can be handled. In [8], Rezaeyan R. and Farnoosh R. present an application of the continuous Kalman-Bucy filter for a RC circuit. The analytic solution of the resulting stochastic integral equations are found using the Ito formula. In [9], Ang examine the use of a simple stochastic differential equation in the modelling of an epidemic. Real data for the Singapore SARS outbreak are used for a detailed study. The model is solved numerically and implemented on matlab, with further analysis and refinement. This article is built around several matlab programs and serves to provide a practical and accessible introduction to numerical methods for a stochastic model for epidemics. In [10], Simha A. et al. model the region-wise trends of the evolution to COVID-19 infections using a stochastic SIR model. The SIR dynamics are expressed using Itô stochastic differential equations. In [11], Zhang et al. consider stochastic mathematical model for COVID-19. Firstly, the formulation of a stochastic susceptible–infected–recovered model is presented. Secondly, they devote with full strength our concentrated attention to sufficient conditions for extinction and persistence. Thirdly, they examine the threshold of the proposed stochastic COVID-19 model, when noise is small or large. Finally, they show the numerical simulations graphically using MATLAB.

Considering the studies mentioned above, the main features of the work finds an appropriate stochastic differential equation model for the given COVID-19 data. So in this research, we have examined by SDEM the COVID-19 data in Turkey between 11.03.2020 and 09.06.2020. In here, parameter values have been estimated from real data using the Maximum Likelihood procedure and approximate solutions of the established stochastic differential equation model have been obtained by using Euler-Maruyama method.

2. Stochastic calculus

2.1. Probability Theory

Stochastic differential equations provide a link between probability theory and the much older and more developed fields of ordinary and partial differential equations [4]. So, the terms of probability theory are significant in stochastic calculus. For instance, a stochastic integral is a random variable and the solution of a stochastic differential equation at any fixed time is a random variable [12]. Therefore, if anybody wants to learn about probability theory in detail, [3, 12-18] can be refered.

2.2. Itô Stochastic Differential Equation

Modeling physical systems by ordinary differential equations (ODEs) ignores stochastic effects. Addition of random elements into the differential equations leads to what is called stochastic differential equations (SDEs), and the term stochastic is called noise. A stochastic differential equation (SDE) is a differential equation in which one or more of the terms is a stochastic process, thus resulting in a solution which is itself a stochastic process [8].
A typical one dimensional Itô stochastic differential equation has the form

$$X(t,\omega) = X(0,\omega) + \int_0^t f\big(s, X(s,\omega)\big)\,ds + \int_0^t g\big(s, X(s,\omega)\big)dW(s,\omega) \tag{2.1}$$

and differential form

$$dX(t) = f(t,X) + g(t,X)dW(t) \tag{2.2}$$

for $0 \le t \le T$ with $X(0,\cdot) \in H_{RV}$, H_{RV} is the Hilbert space of random variables and $X(t,\omega)$ is a stochastic process not a deterministic function. $W(t,\omega) = W(t)$ is a Wiener process of Brownian motion and since it is nowhere differentiable. $W(t),\ t \ge 0$ is a continuous stochastic process with stationary independent increments such that $W(0) = 0$, $\int_c^d dW(s) = W(d) - W(c) \sim N(0, d-c)$ for all $0 \le c \le d$. The function f is often called the drift coefficient of the stochastic differential equation while g is referred to as the diffusion coefficient [12].

3. Material and methods

3.1. Euler-Maruyama Method

An Ito stochastic differential equation on the interval [0,T] has the form

$$dX(t) = f(t, X(t);\ \theta)dt + g(t, X(t);\ \theta)dW(t) \tag{3.1}$$

is considered where $\theta \in \mathbb{R}^m$ is a vector of parameters that are unknown for $0 \le t \le T$ where $X(0) = X_0$, $X_0 \in H_{RV}$ and $W(t)$ is the Wiener process.
As the exact solution to a stochastic differential equation is generally difficult to obtain, it is useful to be able to approximate the solution. Euler-Maruyama method is a simple numerical method. When applied to (3.1), Euler's method has the form

$$X_{j+1,i}^{(m)}\left(t_{i-1}+(j+1)\frac{\Delta t}{K}\right)=X_{j,i}^{(m)}\left(t_{i-1}+j\frac{\Delta t}{K}\right)$$
$$+f\left(t_{i-1}+j\frac{\Delta t}{K},\,X_{j,i}^{(m)}\left(t_{i-1}+j\frac{\Delta t}{K}\right)\right)\frac{\Delta t}{K}+ \qquad (3.2)$$
$$+g\left(t_{i-1}+j\frac{\Delta t}{K},\,X_{j,i}^{(m)}\left(t_{i-1}+j\frac{\Delta t}{K}\right)\right)\sqrt{\frac{\Delta t}{K}}\,\eta_{j,i}^{(m)}$$

for $i=0,1,2,\ldots,N-1$ and $j=0,1,2,\ldots,K-1$ where $\Delta t=\frac{T}{N}$, $t_i=i\Delta t$, and $\Delta t_i=t_{i+1}-t_i=\frac{\Delta T}{K}$, $\eta_{j,i}^{(m)}\sim N(0,\Delta t)$, $\Delta W(t,w)=W(t_{i+1},w)-W(t_i,w), \Delta W(t,w)\sim N(0,\Delta t_i)$, and m indicates a simulation number. This means that by changing number K, m times then m simulations realized, [14].

In this research, the extreme values of approximation trajectories $X_{j+1,i}^{(m)}$ by applying Euler-Maruyama Approximation Method are achieved.

3.2. Parameter Estimation for Stochastic Differential Equations

The problem is to find an estimate of the vector θ given these $N+1$ data points. Two estimation methods are a maximum likelihood estimation method and a nonparametric estimation method, [12].

In this study, we use the maximum likelihood estimation method to estimate parameters from the real data.

3.3. Maximum-Likelihood Estimation Method

Let $p(t_k,x_k|t_{k-1},x_{k-1};\boldsymbol{\theta})$ be the transition probability density of (t_k,x_k) starting from (t_{k-1},x_{k-1}) given the vector $\boldsymbol{\theta}$. Suppose that the density of the initial state is $p_0(x_0|\boldsymbol{\theta})$.

In maximum likelihood estimation of $\boldsymbol{\theta}$, the joint density

$$D(\boldsymbol{\theta})=p_0(x_0|\boldsymbol{\theta})\prod_{k=1}^{N}p(t_k,x_k|t_{k-1},x_{k-1};\boldsymbol{\theta})$$

is maximized over $\boldsymbol{\theta}\in\mathbb{R}^{\mathbf{m}}$ [12]. However, to avoid small numbers on a computer, it is more convenient to minimize the function $L(\boldsymbol{\theta})=-\ln\big(D(\boldsymbol{\theta})\big)$ which has the form

$$L(\boldsymbol{\theta})=-\ln\big(p_0(x_0|\boldsymbol{\theta})\big)-\textstyle\sum_{k=1}^{N}\ln\big(p(t_k,x_k|t_{k-1},x_{k-1};\boldsymbol{\theta})\big).$$

4. A Stochastic Model for a COVID-19 outbreak

A simple stochastic model for a disease such as COVID-19 that spreads very quickly is given by the stochastic differential equation

$$dX(t)=\theta_1X(t)\big(n-X(t)\big)dt+\theta_2X(t)dW(t),\qquad X(0)=X_0,\quad 0\le t\le T \qquad (3.3)$$

where $X(t)$ is the number of infected and susceptible individuals at time t (days), θ_1 and θ_2 are real constants, n is the total number of individuals in that community and $W(t)$ is a random variable representing a standard Wiener process. In here, it is obvious that drift represents the deterministic portion of the model, while diffusion represents the stochastic component [9].

In this study, we use equation (3.3) in the modeling of an infectious disease, using the outbreak of COVID-19 data in Turkey between 11.03.2020 and 09.06.2020.

4. An application

One of the important application areas of stochastic differential equations is population biology such as epidemic model. In this study, the COVID-19 data in Turkey between 11.03.2020 and 09.06.2020 are examined. This data set is accessed with the help of World Health Organization (WHO) shared the link via https://www.who.int/ [1], and the data set is shown in Table 1.

Table 1. *The COVID-19 data in Turkey between 11.03.2020 and 09.06.2020*

Day	Total number of confirmed cases	Day	Total number of confirmed cases	Day	Total number of confirmed cases	Day	Total number of confirmed cases
1	1	24	20921	47	110130	70	151615
2	1	25	23934	48	112261	71	152587
3	5	26	27069	49	114653	72	153548
4	5	27	30217	50	117589	73	154500
5	5	28	34109	51	120204	74	155686
6	47	29	38226	52	122392	75	156827
7	47	30	42282	53	124375	76	157814
8	191	31	47029	54	126045	77	158762
9	191	32	52167	55	127659	78	159797
10	670	33	56956	56	129491	79	160979
11	947	34	61049	57	131744	80	162120
12	1236	35	65111	58	133721	81	163103
13	1529	36	69392	59	135569	82	163942
14	1872	37	74193	60	137115	83	164769
15	2433	38	78546	61	138657	84	165555
16	3629	39	82329	62	139771	85	166422
17	5698	40	86306	63	141475	86	167410
18	7402	41	90980	64	143114	87	168340
19	9271	42	95591	65	144749	88	169218
20	10827	43	98674	66	146457	89	170132
21	13531	44	101790	67	148067	90	171121
22	15679	45	104912	68	149435	91	172114
23	18135	46	107773	69	150593		

We shall use following stages of investigation to solve our problem. Firstly, the parameters of SDE established for the COVID-19 data that mentioned above are estimated by using maximum likelihood estimation method. Then, we have obtained reasonable Stochastic Differential Equation (SDE) based on the given COVID-19 data. Moreover, by applying Euler-Maruyama Approximation Method trajectories of SDE are achieved and they are demonstrated in Figure 1.

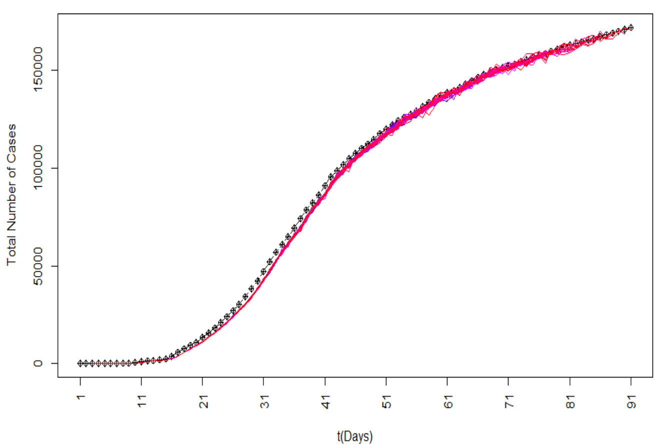

Figure 1. *The COVID-19 data in Turkey between 11.03.2020 and 09.06.2020 (black line) and approximate EM trajectories (colored lines) of the SDE*

Using the maximum likelihood estimation method, $\hat{\theta}_1 = 0.000004$ and $\hat{\theta}_2 = 0.98$ are obtained. If these estimated parameters are considered in the equation (3.3), $dX(t)$ has the following form

$$dX(t) = 0.000004X(t)\big(172114 - X(t)\big)dt + 0.98X(t)dW(t), \quad X(0) = 1 \tag{3.4}$$

Hence, a reasonable stochastic differential equation based on COVID-19 data is achieved. Furthermore, the approximate EM values of the data set $\widehat{X}(t_i)$, $i = 1,2,\ldots,91$ are obtained corresponding to the COVID-19 data in Turkey between 11.03.2020 and 09.06.2020. Approximate trajectories taken randomly according to $\widehat{X}(t_i)$ for $i = 91$ is given in Table 2.

The performances of SDE are established by Chi-Square criteria. The results are acquired by using statistical software R-Studio. These results are also corroborated by graphical representation in Figure 2.

Table 2. *The COVID-19 data in Turkey between 11.03.2020 and 09.06.2020 and the approximate $\hat{X}(t_{91})$ EM values of the data set*

Total number of confirmed cases	EM	Total number of confirmed cases	EM	Total number of confirmed cases	EM	Total number of confirmed cases	EM
1	1.0	20921	20979.3	110130	110413.4	151615	150807.5
1	1.0	23934	24225.2	112261	113325.3	152587	152389.3
5	5.1	27069	27249.7	114653	113652.6	153548	153784.1
5	5.0	30217	29969.5	117589	117343.4	154500	155002.8
5	5.0	34109	34189.8	120204	121549.8	155686	156362.2
47	46.8	38226	38414.4	122392	124744.7	156827	154897.3
47	47.3	42282	42731.3	124375	125322.3	157814	158103.2
191	193.6	47029	47344.9	126045	127087.2	158762	162560.7
191	193.0	52167	52638.2	127659	127618.1	159797	160550. 4
670	684.4	56956	56883.4	129491	131721.1	160979	164169.5
947	944.5	61049	61073.1	131744	131106.9	162120	162024.3
1236	1266.3	65111	66423.3	133721	134270.7	163103	163335.7
1529	1526.7	69392	70377.5	135569	135853.1	163942	164039.0
1872	1893.8	74193	74002.2	137115	133322.9	164769	163312.3
2433	2451.1	78546	80126.2	138657	137647.4	165555	162744.7
3629	3730.0	82329	84909.8	139771	138586.6	166422	167981.2
5698	5693.9	86306	87982.4	141475	139412.4	167410	166266.3
7402	7416.3	90980	91190.2	143114	144217.7	168340	166437.7
9271	92478.0	95591	95935.5	144749	147099.5	169218	169196.8
10827	10909.2	98674	100407.8	146457	147548.3	170132	171263.2
13531	13660.9	101790	101573.0	148067	146751.3	171121	170532.4
15679	15759.2	104912	106494.3	149435	149706.2	172114	172661.0
18135	18290.2	107773	107175.1	150593	149324.4		

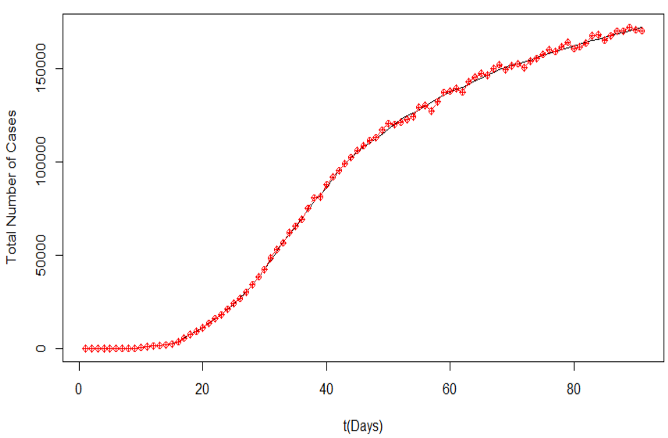

Figure 2. *The COVID-19 data in Turkey between 11.03.2020 and 09.06.2020 (black line) and its EM trajectory starting from $\hat{X}(t_{91})$ (red line)*

5. Conclusion

In this study COVID-19 data in Turkey between 11.03.2020 and 09.06.2020 is investigated by Stochastic Differential Equation Modeling (SDEM). Firstly, parameters of SDE established for the COVID-19 data that mentioned above are estimated estimated $\hat{\theta}_1 = 0.000004$ and $\hat{\theta}_2 = 0.98$ by using maximum likelihood estimation method. Then, we have obtained reasonable Stochastic Differential Equation (SDE) based on the given data. Moreover, by applying Euler-Maruyama Approximation Method trajectories of SDE are achieved. The acceptancy of equation (3.3) with parameters θ_1 and θ_2 to consider in our study for COVID-19 data in Turkey is provided with Chi-square value equal to 7735 of Goodness of fit test. The performances of trajectories according to $\hat{X}(t_{91})$ is established by Chi-Square criteria. The acceptancy of Chi-Square criteria for $\hat{X}(t_{91})$ is realized by values of $p_{value} = 0.2451$ and we do not reject the null hypothesis that given COVID-19 data is fit to last values of EM estimations. As a result, the SDEM we established fits well with COVID-19 data Turkey between 11.03.2020 and 09.06.2020.

References

[1]. World Health Organization (WHO). Coronavirus. Available from: https://www.who.int/health-topics/coronavirus#tab=tab_1 (Accessed: April 15, 2021).

[2]. Iacus S.M., Simulation and Inference for Stochastic Differential Equations with R Examples. USA: Springer, 2008.

[3]. Oksendal B., Stochastic Differential Equations an Introduction with Applications, 5th ed., Corrected Printing. New York: Springer-Verlag Heidelberg, 2003.

[4]. Kostrista E., Çibuku D., "Introduction to Stochastic Differential Equations", Journal of Natural Sciences and Mathematics of UT, 3, (2018), 5-6.

[5]. Ince N., Shamilov A., "An application of new method to obtain probability density function of solution of stochastic differential equations", Applied Mathematics and Nonlinear Sciences, 5.1 (2020), 337-348.

[6]. Mahrouf M. et al. "Modeling and forecasting of COVID-19 spreading by delayed stochastic differential equations", Axioms 10.1, (2021), 18.

[7]. Bak J., Nielsen A. and Madsen H., "Goodness of fit of stochastic differential equations", 21th Symposium I Anvendt Statistik, Copenhagen Business School, Copenhagen, Denmark. 1999.

[8]. Rezaeyan R., Farnoosh R., "Stochastic Differential Equations and Application of the Kalman-Bucy Filter in the Modeling of RC Circuit", Applied Mathematical Sciences, 4, (2010), 1119-1127.

[9]. Ang K. C., "A Simple Stochastic Model for an Epidemic-Numerical Experiments with MATLAB", The Electronic Journal of Mathematics and Technology, 1, (2007), 117-128.

[10]. Simha A., Prasad R.V., and Narayana A.,. "A simple stochastic sir model for covid 19 infection dynamics for karnataka: Learning from europe." arXiv preprint arXiv:2003.11920, (2020).

[11]. Zhang Z., et al. "Dynamics of COVID-19 mathematical model with stochastic perturbation." Advances in Difference Equations, 2020.1, (2020), 1-12.

[12]. Allen E., Modeling with Itô Stochastic Differential Equations. USA: Springer, 2007.

[13]. Shamilov A., Measurement Theory, Probability and Lebesgue Integral, Eskişehir: Anadolu University Publications, 2007.

[14]. Shamilov A., Differential Equations with Theory and Solved Problems, Turkey: Nobel Publishing House, 2012.

[15]. Shamilov A., Probability Theory with Conceptional Interpretations and Applications. Turkey: Nobel Publishing House, 2014.

[16]. Andersson H., Britton T., Stochastic Epidemic Models and Their Statistical Analysis. New York: Springer, 2000.

[17]. Allen L.J.S., An Introduction to Stochastic Processes with Applications to Biology. Upper Saddle River, New Jersey: Pearson Education Inc., 2003.

[18]. Higham D.J., "An Algorithmic Introduction to Numerical Simulation of Stochastic Differential Equations", SIAM Review, 43, (2001), 525-546.