

# Modelling Covid-19 Deaths in Ghana as a Discrete State Process in Continuous Time

Osei Antwi<sup>1,\*</sup>, Abdul Martinu Issah<sup>2</sup>

<sup>1</sup>Department of Mathematics & Statistics, Accra Technical University, Accra, Ghana

<sup>2</sup>Research Department, Fair Wages & Salaries Commission, Accra, Ghana

## Email address:

oseiantwi@yahoo.co.uk (O. Antwi), abdulmartinuissah@gmail.com (A. M. Issah)

\*Corresponding author

## To cite this article:

Osei Antwi, Abdul Martinu Issah. Modelling Covid-19 Deaths in Ghana as a Discrete State Process in Continuous Time. *American Journal of Applied Mathematics*. Vol. 8, No. 6, 2020, pp. 344-355. doi: 10.11648/j.ajam.20200806.17

**Received:** November 30, 2020; **Accepted:** December 18, 2020; **Published:** December 31, 2020

---

**Abstract:** We propose a stochastic process modelling of covid-19 deaths in Ghana. The objective is to accurately capture the death processes resulting from the pandemic and to predict future deaths resulting from Covid-19 infections in Ghana. The mathematical derivation is based strictly on the compound Poisson process, a class of a Levy process. The model is verified by using empirical data of deaths resulting from Covid-19 from the onset of the pandemic up to the time of writing this report. That is, Covid-19 deaths in Ghana from March to August 2020. The method departs slightly from the usual differential equations used in modeling pandemics due to the unique occurrence of deaths from the disease in Ghana. As the methods are basically compound Poisson process, we delve into Levy processes as it allows us to effectively simulate the future behaviour of the death process. To test the effectiveness of the model, we compared the simulated results to the actual reported number of deaths from Covid-19 cases in Ghana from March to August 2020. The results show that at a 95% confidence interval there is no significant difference between the actual deaths and the simulated results. The results of the simulation, when extended to February 2021 (one year after the advent of the pandemic) shows that if the current conditions remain same, that is, if there is no immediate intervention by the discovery of an effective drug or a vaccine, then the number of deaths could reach four hundred and forty six (446) by February 28, 2020.

**Keywords:** Death Event, Death Event Sizes, Poisson Process, Compound Poisson Process, Levy Process

---

## 1. Introduction

As is widely known by now, the first case of Corona Virus was reported in Wuhun, a city in China in late December 2019. In February 2020, Iran, Italy and Spain reported the presence of the virus in their territories. Events then unfolded quickly and by March 2020, The World Health Organization [1] had declared the situation a pandemic, the first since the Spanish Flu in 1918. Figure 1 shows The Guardian [2] World map of coronavirus infections as at the end of August 2020.

Ghana reported its first case of corona virus in Accra on 12<sup>th</sup> March, 2020 and the disease has since spread to all other regions in the country, infecting over 40,000 people by end of August 2020.

Although a large proportion of people infected with the virus show no symptoms or may only develop mild

symptoms such as increased body temperature, mild cough, loss of taste and smell and do not develop serious illness, older people and those already carrying some kind of underlying medical conditions such as diabetes, chronic respiratory disease, or cancer may develop serious illness.

The outbreak of the pandemic has been devastating to the world. As it raged, it has destroyed many economies, killed almost a million people by August 2020, and has led to lockdowns in many countries and regions. Orlik T. et al., [3] estimates that the cost of Covid-19 to the world economy could top \$2.7 trillion, equivalent to United Kingdom's economy. Developing countries, Western economies and emerging markets such as India and Brazil have been significantly impacted. Developing countries for instance, have had to do without support from western governments who are themselves reeling from the impact of the pandemic. Consequently, several forms of development assistance, debt

relief & debt restructuring have been suspended. Despite multilateral institutions like IMF and the World Bank providing a lifeline for African countries to access funds to

combat the virus many countries including Ghana have cut rates and announced various monetary policies.



*Figure 1. Corona virus infections around the world - August 2020.*

Although there has been multifaceted response from governments and multilateral institutions toward mitigating the economic impact, any gains cannot be consolidated without a meaningful model of the future trend of the disease especially in terms of loss of human capacity. It is in this regard that we have decided to investigate and develop a mathematical model to predict Covid-19 fatalities in Ghana. To this end, we track the pattern of deaths resulting from the virus and use stochastic models predict deaths resulting from Covid-19 in the immediate future in Ghana.

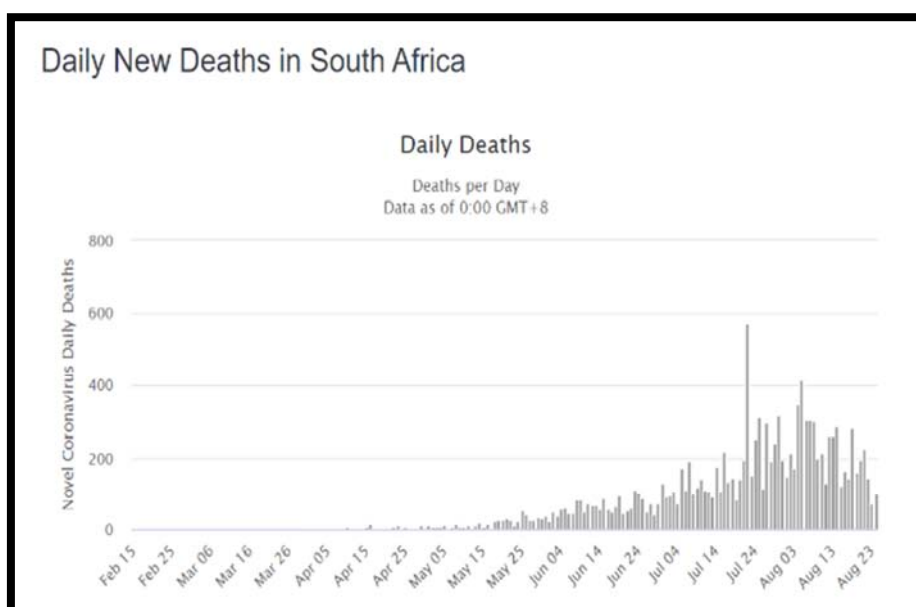
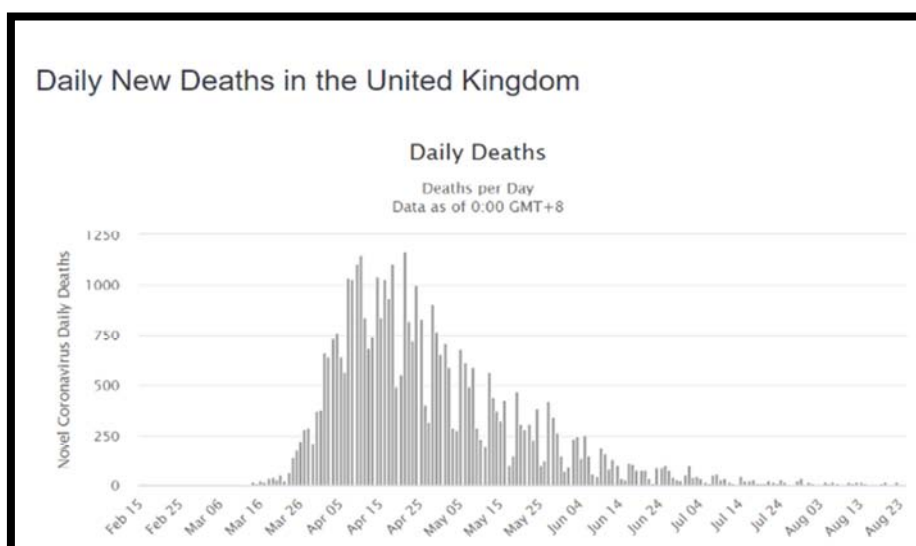
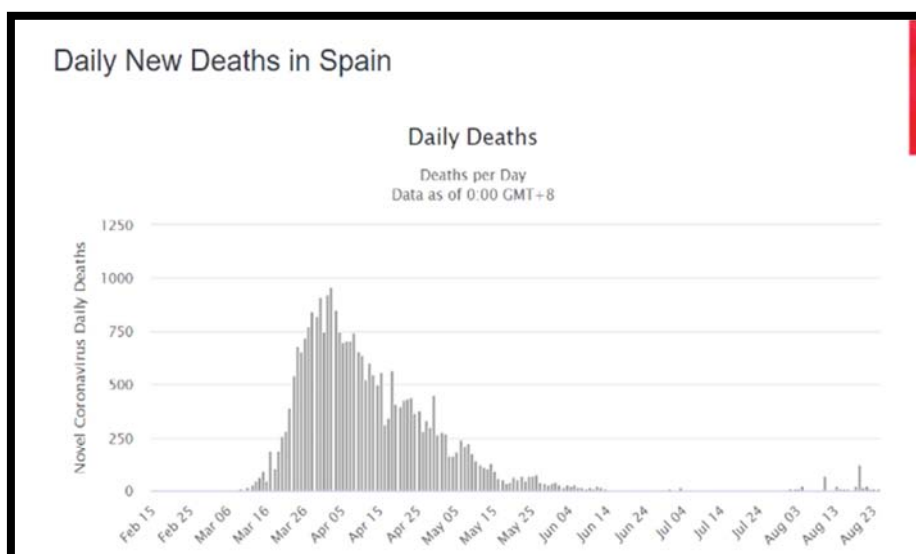
The study is based on a stochastic model, specifically, the compound Poisson process, a form of a Lévy process. As it's synonymous with most mathematical models, the predictive ability of the model is limited by the reliability of the available empirical data and in addition, the model is subject to several assumptions and thus do not convey direct clinical information and caution to the general public. Castorina P, et al., Dehkordi A.H. et al., [4, 5]. However, we believe that the model's highlights will be of great benefit and offer reliable information to enable for a quick assessment of the severity of the pandemic and help guide government and health officials in defining or developing national and regional strategies to combat the disease. In addition, the paper also provides mathematical insights and a rich theoretical framework to investigate the dynamics of deaths from the pandemic.

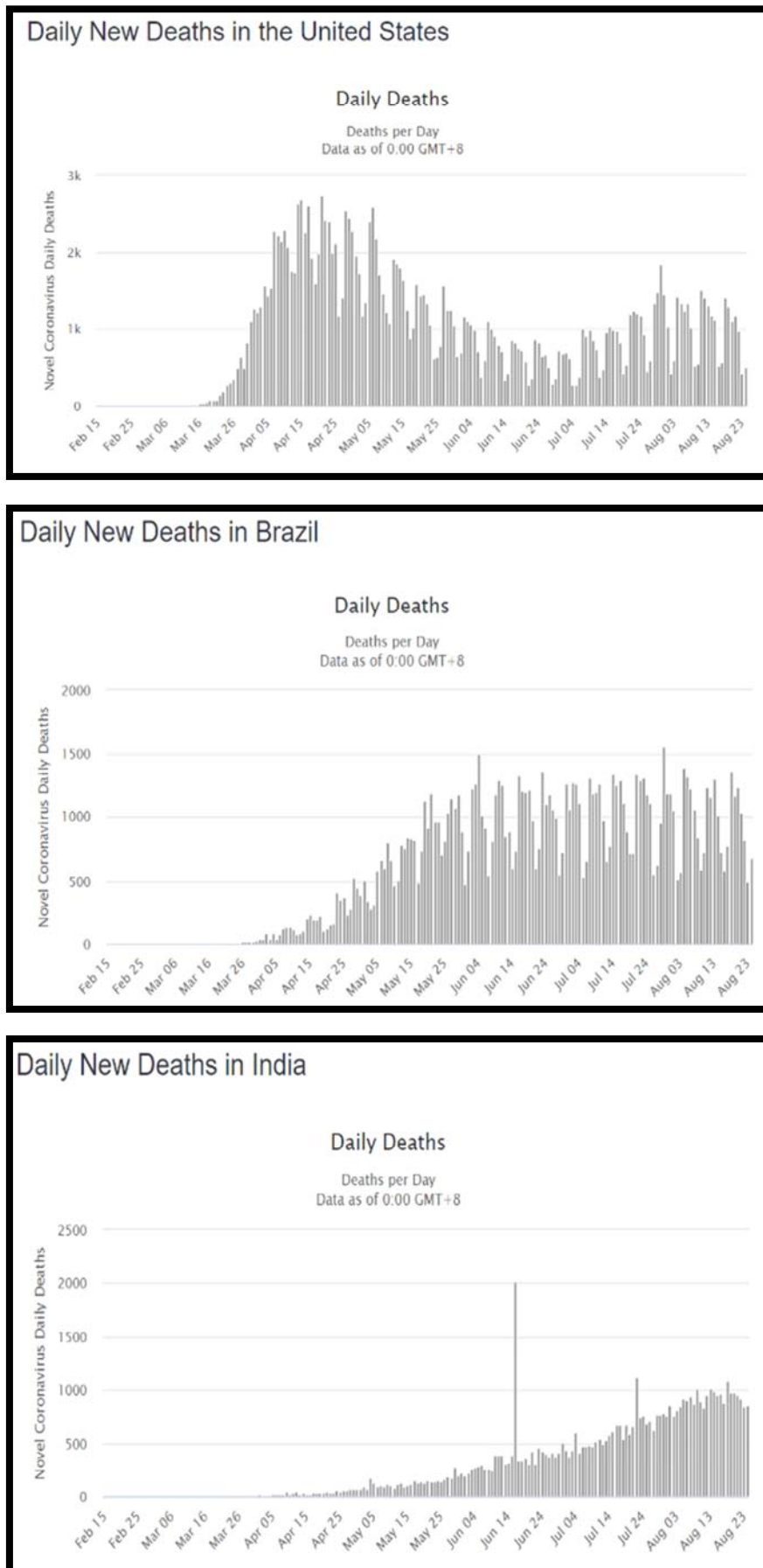
It must be stated here that we are more interested in the number of deaths because it gives a more direct or accurate measure of the impact of the virus as compared to the number of infected cases. This is because in the first place many carriers of the virus do not show any symptoms (asymptomatic) or may develop only mild symptoms which in most cases does not lead to hospitalisation. Secondly, the number of tests carried out in Ghana was only meant for

targeted sample of the population through contact tracing, routine surveillance and other related activities. Thus, the population was not mass tested and a such a large proportion of Covid-19 infections were not detected. The number of reported cases is in effect, was only an unsatisfactory representation for the total infections in the populace Li R. et al., [6]. The true reflection of the virus situation in the country is thus difficult to estimate from the reported infections. On the other hand, the number of deaths from Covid-19 related diseases is a more accurate measure of the advance of the epidemic and its prevalence as deaths from the disease can hardly go undetected. Although the number of deaths attributed to Covid-19 from official sources may have some inaccuracies and uncertainties of their own, primarily due to individual government's recording and reporting protocols, nevertheless, it is still realistic to assume that the evolution of the number of confirmed deaths bears a direct relation to the dynamics of the impact of the disease. Under these circumstances, coupled with the absence of other reliable sources for the estimates of the number of infected cases, we decided to model death dynamics rather than infection rates. Thus, the main objective is to model the daily deaths from covid-19 as well as the cumulative number of deaths over time. Furthermore, the paper makes predictions of daily deaths and the cumulative deaths in the immediate future.

Figure 2 shows the graph of Covid-19 deaths in some countries in North and South America, Europe, Asia and Africa. The graphs indicate that there are stark differences in the way death cases have evolved in the various countries. It is realised that whereas in other countries deaths are daily events and occur in hundreds or thousands, death cases in Ghana do not occur daily but rather occur over irregular time

intervals and are usually in single or double figures.





**Figure 2.** Covid -19 Deaths around the World: March–August 2020.

Based on the empirical evidence it will be inappropriate to model Covid-19 deaths in Ghana using existing models proposed in other countries. It is therefore imperative to develop models suitable to Ghana's unique situation in order to capture the true picture of losses from the pandemic in the country. The aim of this paper is therefore to develop an appropriate mathematical model that can suitably predict daily death toll from covid-19 and in addition provide a measure of the cumulative number of deaths in the near future.

Mathematical models of the evolution of the number of infections in a pandemic is mostly dominated by ordinary differential equations resulting in exponential growth of infections. The most general and frequently used framework is the Susceptible-Infectious-Recovered (SIR) class of models and its variants. Piccolomini, E. L., & Zama, F.; J. Kucharski et al.; Berger, K. H., & Mongey, S.; Walker P. G. T., et al., [7–10]. The basic and the most useful of the SIR models is the Richards growth model defined by ordinary differential equations Richards F.J.; Hsieh Y-H., [11, 12]. Mathematical models using this approach forecast Covid-19 deaths by estimating the number of deaths as a percentage of the number of infectious cases Ivorra B, et al.; Sameni R.; Crokidakis N.; Bastos S. B. & Cajueiro, D. O.; Andersson H. & Britton, T., [13–17].

Other models directly compute the daily deaths. One example of this approach is by Roman Cherniha, et al., [18] who proposed a model based on non-linear ordinary differential equations. The results of the model were shown to be in agreement with measured public data in China and Austria and was able to make predictions with less than 10 percent error. Pham, H., [19] used ordinary differential equations to develop a model to estimate the cumulative number of deaths in the United States. The model's results were compared to two related existing models based on a selection criterion including SSE (Sum of Squared Error), MSE (Mean Squared Error). The results show that the proposed model fits significantly better than the other two related models. The model predicted with a 95% confidence that the expected total death toll will be between 60,951 and 63,249 deaths by 22 April 2020. Pham, H., [20] used partial differential equations to predict the number of Covid-19 deaths in the United States. The model was based on a set of indicators such as recovered cases, daily new cases, total cases as well as the recovery rate. The modelling results were found to be in agreement with real time United States daily deaths as well as in predicting the number of deaths in Italy and The United Kingdom. Another study, M. Yousaf, et al., [21] analysed data from Islamabad and produced a forecast of Covid-19 confirmed cases as well as the number of deaths and recoveries in Pakistan using the Auto-Regressive Integrated Moving Average Model (ARIMA). The model when fitted to the live data revealed that the cumulative number of Covid-19 confirmed cases, deaths and recoveries show exponential growth over time in regards to number of confirmed cases in Pakistan. The ARIMA model also had

higher fitting and forecasting accuracy than exponential smoothing and captures both the seasonal and non-seasonal forecasting trend.

## 2. Methodology

### 2.1. Existing Models to Estimate Covid-19 Deaths

On the models that directly predict Covid-19 deaths, Roman Cherniha & Vasyly Davydovych proposed a non-linear ordinary differential equations of the form

$$\frac{du}{dt} = u(a - bu), u(0) = u_0 \geq 0, \quad (1)$$

where  $u(t)$  predicts the total number of the Covid-19 cases and  $a$  and  $b$  are positive constants.

By setting  $u = v + w$  we have

$$\frac{dv}{dt} = k(t)u \quad v(0) = v_0 \geq 0 \quad (2)$$

where  $w$  are recoveries and  $v$  represents deaths at time  $t$  and  $k(t) > 0$  reflects the effectiveness of the health care system of the country or region.

Equations (1) and (2) are generalized as

$$\frac{du}{dt} = (u - v)(a - b(u - v)\gamma) \quad (3)$$

$$\frac{dv}{dt} = k(t)(u - v) \quad (4)$$

$$\frac{dw}{dt} = w(a - k(t) - bw\gamma) \quad (5)$$

$$u(0) = u_0 \geq 0, v(0) = v_0 \geq 0, w(0) = w_0 = u_0$$

Solving Equations (3), (4), (5) and making substitutions the integral representation is giving by

$$v(t) = v_0 + w(t) + \int_0^t k(\tau)w(\tau) d\tau$$

Setting  $k(t) = k_0 e^{-\alpha t}$ ,  $\alpha > 0$ , the exact solution of equations (1) and (2) are

$$u(t) = \frac{au_0 e^{at}}{a + bu_0(e^{at} - 1)}$$

$$v(t) = ak_0 u_0 \int_0^t \frac{e^{(a-\alpha)\tau}}{a + bu_0(e^{a\tau} - 1)} d\tau + v_0$$

When the model was applied to data on the outbreak in China the number of deaths was obtained as

$$v|_{t=T} \equiv 571k_0 \int_0^T \left( \frac{e^{0.21\tau}}{1 + \frac{571}{80,000}(e^{0.28\tau} - 1)} \right) d\tau + 17$$

Hoang Pham also used partial differential equations to predict Covid-19 deaths in the United States. The generalized mathematical model follows the Partial Differential Equation

$$\frac{\partial p(t,x)}{\partial t} \alpha [I_{(t,x)}] b(t,x) p(t,x) (a(t,x) - p(t,x) + (1 - \alpha [I_{(t,x)}]) b(t,x) (a(t,x) - p(t,x))) \quad (6)$$

where

$a(t,x)$  = the total number of deaths

$b(t,x)$  = death rate per person

$p(t,x)$  = cumulative number of deaths by time  $t$

$x$  = a set of indicators

$\alpha [I_{(t,x)}] \in (0,1)$  = recovery rate by time  $t$

Solving the partial differential equation in Equation (6) produces two solutions  $p_1(t)$  and  $p_2(t)$  estimating the cumulative number of deaths

$$p_1(t) = \frac{\alpha}{1 + \left( \frac{c}{\beta + e^{bt}} \right)}$$

when  $\alpha [I_{(t,x)}] = 0$

$$p_2(t) = \frac{\alpha(1 + dt)}{1 + \left( \frac{c}{\beta + e^{bt}} \right)}$$

where  $a, b, c, d$  and  $\beta$  are the unknown constants. The daily death toll is obtained as

$$r_2(t) = \frac{\alpha d(\beta + e^{bt})}{(\beta + e^{bt} + c)} + \frac{(1 + dt)(bce^{bt})}{[\beta + e^{bt} + c]^2}$$

## 2.2. Modelling Ghana Covid-19 Deaths

Consider Figure 3 which shows the plot of Covid-19 deaths in Ghana from March–August 2020. We realise that deaths occur over infrequent days and when they occur, the number of deaths also differ in sizes on the reporting day. We will refer to these reporting days as death events. So, for instance, there were 7 death events in April but May reported 10 death events. In addition, at each death event the number of deaths vary in size. Thus, for instance, on the first death event on 23rd March had death size 1. However, on the ninth death event on April 28 the deaths size was 5.

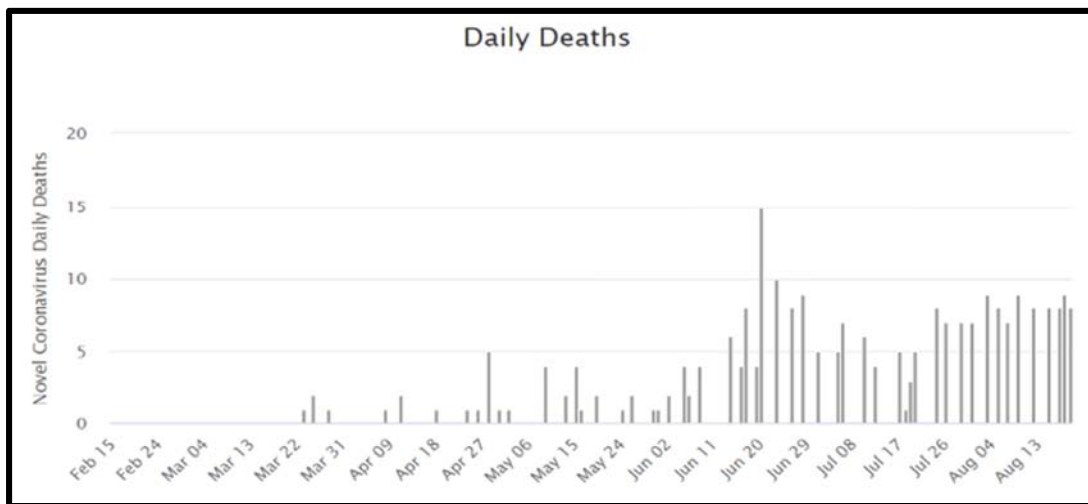


Figure 3. Deaths events from Covid -19 in Ghana. March–August 2020.

The process can be modelled as follows:

- 1) Let  $N(t)$  counts the number of deaths events and  $T_n, n = 1, 2, \dots$  be the random times at which the death events occur. At time  $t = 0$ , that is, at the beginning, before any death has occurred, the number of death events was zero. That is, at  $t = 0, N(t) = 0$  and we write  $N(0) = 0$ .  $N(t)$  stays at the level 0 until some random time  $T_1$  when the first death event occurs and  $N(t)$  counts the death events as 1 so we write  $N(T_1) = 1$ . We do not observe any death event until another random time  $T_2$  when the next death event occurs and we have  $N(T_2) = 2$ .
- 2) It is observed that the number of deaths can only add up and thus given any two times  $s$  and  $t$  such that  $s < t$ , the number of deaths at  $s$  is always less than the number of deaths at  $t$  and write  $N(s) < N(t)$  for  $s < t$ .  $N(t)$  can thus be considered as a non-decreasing function.
- 3) The increment of death events  $N(t) - N(s)$  is

independent of the previous information. Thus if  $(\Omega, \mathcal{F}, \mathbb{P})$  is a probability space admitting  $N(t)$  and the  $\sigma$ -algebra  $\mathcal{F}_t$  such that  $\mathcal{F}_t = \sigma\{N(s): 0 \leq s \leq t\}$  is a filtration satisfying  $N(t) \in \mathcal{F}_t$  for all  $t$ , then  $N(t) - N(s)$  is independent of  $\mathcal{F}_s$ .

The characteristics of the death event process can therefore be summarised as follows:

- i.  $N(0) = 0$ .
- ii.  $N(t)$  counts the number of death events and it is always an integer valued function.
- iii. For all  $s < t$ ,  $N(s) < N(t)$ . That is,  $N(t)$  is a non decreasing function
- iv. The increment of death events  $N(t) - N(s)$  is independent of  $\mathcal{F}_s$ ,
- v.  $N(t)$  is a discrete state process occurring in continuous time frame

It happens that characteristics  $(i - v)$  above of the death event process satisfies the criteria for  $N(t)$  to be approximated by the Poisson process. If we define a constant  $\lambda$  as the mean number of death events, then for all  $s < t$ ,  $N(t) - N(s)$  is a Poisson random variable with parameter  $\lambda(t - s)$ .

#### Definition

Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a probability space with a given filtration  $\mathbb{F} = \{\mathcal{F}_t\}_{t \geq 0}$  and let  $\lambda$  be a nonnegative real number. The process  $N(t)$  which counts the number of death events is a Poisson process with intensity  $\lambda$  with respect to the filtration  $\mathbb{F}$  and satisfies the following conditions.

- i.  $N(0) = 0$
- ii. For all  $s < t$ ,  $N(t) - N(s)$  is independent of  $\mathcal{F}_s$
- iii. For all  $s < t$ , the conditional distribution of the increment  $N(t) - N(s)$  is given by

$$\mathbb{P}\left[(N(t) - N(s)) \mid \mathcal{F}_t\right] = \frac{e^{-\lambda(t-s)} (\lambda(t-s))^n}{n!},$$

for all  $n = 0, 1, 2, \dots$  (7)

#### Probability Distribution of $N(t)$

Given that  $N(t)$  is Poisson and at time  $t$  the number of deaths is  $n$ , then in a small-time interval  $h$ , there will be three possibilities:

1. The probability that by the time  $t + h$ , the number of deaths has increased by 1 to  $n + 1$ , is given by

$$\mathbb{P}[N(t + h) = n + 1 \mid N(t) = n]$$

If the Poisson process has intensity  $\lambda$  then

$$\mathbb{P}[N(t + h) = n + 1 \mid N(t) = n] = \lambda h + o(h)$$

where  $o(h)$  is a small adjustment such that

$$\lim_{h \rightarrow 0} \frac{f(h)}{h} = 0$$

2. As  $h \rightarrow 0$ , there can only be a small number of adjustments in the number of death events. That is, in the minute interval the number of death events increase by 1. This means that

$$\mathbb{P}[N(t + h) > n + 1 \mid N(t) = n] = o(h)$$

3. The probability that the deaths will be greater than  $n + 1$  given that  $N(t) = n$  is almost negligible. Hence

$$\mathbb{P}[N(t + h) = n + 1 \mid N(t) = n] = 1 - \lambda h + o(h)$$

The probability distribution of the increment of the process can be summarised as

$$\mathbb{P}[N(t + h) - N(t) = n] = \begin{cases} \lambda h + o(h) & n = 1 \\ o(h), & n \geq 2 \\ 1 - \lambda h + o(h), & n = 0 \end{cases}$$

#### 1. Distribution of the process up to the first death event

Let's assume that there are no deaths up to time  $t$  and let  $T_1$  be the time to first death event. The time to first death event can be designated as  $T_1 > t$ . But this is the same as the time up to  $t = 0$  or the time to the first death event. The probability of this death event is given as

$$\mathbb{P}(T_1 > t) = \mathbb{P}[N(t) = 0]$$

But  $N(t)$  is Poisson and hence

$$\mathbb{P}[N(t) = 0] = \frac{e^{-\lambda t} (\lambda t)^0}{0!} = \frac{e^{-\lambda t} \cdot 1}{1} = e^{-\lambda t}.$$

The probability that a death event will occur between 0 and  $t$  is thus given by

$$\mathbb{P}(T_1 \leq t) = 1 - e^{-\lambda t}$$

But  $1 - e^{-\lambda t}$  is the distribution function for the exponential distribution. Hence the distribution of the process up to the time of the first death event follows the exponential distribution with parameter  $\lambda$ .

#### 2. Distribution of the process between two death events

By the same argument as in 1 above we can write that in a small-time interval  $h$

$$\mathbb{P}[N(t + h) - N(t) = n \mid N(t) = n]$$

This means that there is no increase in number of deaths and so

$$\mathbb{P}[N(t + h) - N(t) = 0 \mid N(t) = n]$$

This confirms that the increment  $N(t + h) - N(t)$  is independent of the time period  $N(t) = n$

$$\begin{aligned} \mathbb{P}[N(t + h) - N(t) = 0 \mid N(t) = n] \\ &= \mathbb{P}[N(t + h) - N(t) = 0] \\ &= \mathbb{P}[N(t) = 0] \\ &= e^{-\lambda t} \end{aligned}$$

Thus, between two death events the process is distributed exponentially with parameter  $\lambda$ .

#### 2.3. The Infinitesimal Characteristics of Covid-19 Death Events

We realise that deaths in America and elsewhere occur daily in hundreds or thousands. Thus, although individual deaths are discrete the entire death process can be modelled as a continuous process without any loss of generality. This is not the case in Ghana where death events are infrequent and unpredictable making each death event a non anticipatory activity. This allows the introduction of stochastic processes. We notice that approaching a death event from the left side, the limiting value is the value of the process just before a death event. On the other hand, approaching a death event from the



right-hand side, the process has a limit but it is continuous because the process attains the limiting value since the death event had occurred. So, the process is right continuous with left limits. Thus, the process comes from a family of “*continue adroit limit á gauge*” (right continuous with left limits) *cadlag* functions. Essentially, the *cadlag* specification means that the death events are non – anticipative. One could alternatively define the death event process as left continuous with right limits (*caglag*), but then the death events would be anticipative or predictable process as in the case of America and elsewhere. For extensive literature on *cadlag* and *caglag* functions see Cont R., & Tankov P. [22].

Now approaching a death event from the left, if  $N(t)$  is a Poisson process then just before a potential death event  $N(t)$  is written as  $N(t-)$  and defined by

$$N(t-) = \lim_{s \uparrow t} N(s)$$

In a continuous setting the increment process is defined as  $dN(t) = N(t) - N(t-) = N(t) - N(t-dt)$  and takes only two possible values:

1. where no death event occurs and  $dN(t) = 0$ ;
2. where a death event occurs and  $dN(t) = 1$ .

In particular,

$$N(t) = \sum_{n=1}^{\infty} \mathbb{1}_{T_{(n,\infty)}}(t)$$

where

$$\mathbb{1}_{T_{(n,\infty)}}(t) = \begin{cases} 1 & \text{if } t \geq T_n \\ 0 & \text{if } 0 \leq t < T_n \end{cases}$$

## 2.4. The Compound Poisson Process

Thus far, we know that the death events follow the Poisson distribution with intensity  $\lambda$ . But we realise that in addition to the death events the number of deaths or death sizes in each event must also be considered. For example, on the first death event on 23 March 2020, the number of individual deaths was 1 but on the ninth death event on 29 April 2020, the number of individual deaths were nine (9). Consequently, to develop a complete and realistic model to describe the death processes the Poisson process must be extended.

Let's assume that the total number of deaths is a random variable with some probability distribution, identically and independently distributed with finite expectation and variance. Let  $Y(t)$  represents the total number of deaths after time  $t$  and let  $Z(1), Z(2), \dots$  denote the i.i.d sequential death event sizes, then the total number of deaths by time  $t$  is given by the compound Poisson process

$$\mathbb{E}[\exp((Y(t) - Y(s)))] = \mathbb{E}\left[\exp\left(\sum_{k=N(s)+1}^{N(t)} Z(k)\right)\right] = \mathbb{E}\left[\exp\left(\sum_{k=N(s)+1}^{N(t)} Z(k)\right)\right] = \mathbb{E}\left[\exp\left(\sum_{k=1}^{N(t)-N(s)} Z(k)\right)\right]$$

$$Y(t) = \sum_{k=1}^{N(t)} Z(k)$$

where  $N(t)$  is a Poisson process.

### 2.4.1. Levy Representation

Now let's examine the process under levy measures. We diverge into Levy processes as they admit a fair amount of tractability and it is also possible to simulate at a fixed set of dates. Let

$$Y(t) = \sum_{k=1}^{N(t)} Z(k)$$

be a compound Poisson process. The following proposition allows the compound Poisson process to be modelled as a Levy process.

#### Proposition

A Poisson process  $\{Y(t)\}_{t \geq 0}$  is a compound Poisson process if and only if it is a Levy process and its sample paths are piecewise constant functions.

Thus, from proposition 1, the compound Poisson process can be modelled as a Levy process.

### 2.4.2. Expected Number of Deaths

Let  $(Y(t))_{t \geq 0}$  be a compound Poisson process with intensity  $\lambda$  and death event size distribution  $f(x)$ , then  $Y(t)_{t \geq 0}$  is a Levy process with density

$$\lambda = \int_0^T v(dx)$$

$v$  is the Levy measure and represents the expected number per unit time of the size of a death event. We assume  $\int_0^T v(dx) < \infty$  and as such the Levy process has a finite expected number of death events in the interval  $[0, T]$ . In particular, the process has the characteristic triplet  $(a, 0, v)$  and  $v(dx) = \lambda f(dx)dx$  is the arrival rate of death events. We assume the aggregate death events is finite. That is

$$\int_0^T v(dx) < \infty$$

This means that there are finite number of death events within any finite time interval  $[0, T]$ .

Having introduced the process in a Levy framework, we deduce the expected deaths in a given time  $t$  as follows: Given that  $\{N(T) = n\}$ , the death event sizes  $(Y(t))_{t \in \mathbb{R}^+}$  on  $[0, T]$  are independent random variables which are distributed according to  $v(dx)$ . Thus, for any  $s < t \in [0, T]$  the expected increment in  $Y(t)$  is given by



$$\begin{aligned}
&= \mathbb{E} \left[ \exp \left( \sum_{k=1}^{N(t)-N(s)} Z(k) \right) \middle| N(t) - N(s) = n \right] \mathbb{P}(N(t) - N(s) = n) \\
&= e^{-\lambda(t-s)} \sum_{n=0}^{\infty} \frac{\lambda^n (t-s)^n}{n!} \left( \sum_{k=1}^n Z(k) \right) = e^{-\lambda(t-s)} \sum_{n=0}^{\infty} \frac{\lambda^n (t-s)^n}{n!} \prod_{k=1}^n \mathbb{E}[\exp(Z(k))] \\
&= e^{-\lambda(t-s)} \sum_{n=0}^{\infty} \frac{\lambda^n (t-s)^n}{n!} \{\mathbb{E}[\exp(Z(k))]\}^n = \exp\{\lambda(t-s)(\mathbb{E}[\exp(Z(k))]-1)\} \\
&= \exp \left\{ \lambda(t-s) \int_{-\infty}^{\infty} e^{\alpha y} v(dy) - \lambda(t-s) \int_{-\infty}^{\infty} v(dy) \right\} = \exp \left\{ \lambda(t-s) \int_{-\infty}^{\infty} (e^{\alpha y} - 1) v(dy) \right\} \\
&\mathbb{E}[\exp((Y(t) - Y(s)))] = \exp \left\{ \lambda(t-s) \int_{-\infty}^{\infty} (e^{\alpha y} - 1) v(dy) \right\} \\
&\mathbb{E}[\exp((Y(t) - Y(s)))] = \exp \left\{ \lambda(t-s) \int_{-\infty}^{\infty} (e^{\alpha y} - 1) \lambda f(dx) dt \right\}
\end{aligned}$$

the expectation of  $Y(t)$  is computed as the product of the mean number of deaths event times  $\mathbb{E}[N(t)] = \lambda t$  and the mean death event size  $\mathbb{E}[Z]$ , i.e.,

$$\mathbb{E}[Y(t)] = \frac{\partial}{\partial \alpha} \mathbb{E}[e^{\alpha Y(t)}] = \lambda t \int_{-\infty}^{\infty} y v(dy) = \mathbb{E}[N(t)] \mathbb{E}[Z] = \lambda t \mathbb{E}[Z]$$

Variance of  $Y(t)$  is obtained similarly as

$$\text{Var}[Y(t)] = \lambda t \int_{-\infty}^{\infty} y^2 v(dy) = \lambda t \mathbb{E}[|Z|^2] = \mathbb{E}[N(t)] \mathbb{E}[|Z|^2]$$

### 3. Simulating Expected Number of Deaths

The data used in this study was obtained from the database made publicly available by Worldometer [23]. The website lists in a tabular form daily updates of Total Number of Cases (Cumulative), Total Number of New Cases, Total Number of Deaths (Cumulative), Total Number of New Deaths, Recovered Cases, Active Cases, Serious or Critical Cases, Total Number of Cases Per Million Population, Total Number of Deaths Per Million Population, Total Number of Test Per Million Population of every country affected by Corona

Virus. The daily deaths are used as input parameters and we compute the mean number of death events  $\lambda$ . The simulation proceeds by first simulating  $N(t)$  as a Poisson distribution with parameter  $\lambda$ , that is,  $N(t) \sim Po(\lambda)$ . The time to the next death event follows the exponential distribution,  $N(t) \sim \exp(\lambda)$ . The death event sizes are sampled from the normal distribution with mean  $\mu$  and variance  $\sigma^2$ , that is,  $Z(n) \sim \mathcal{Z}(\mu, \sigma^2)$ . This assumption is realistic given the monthly distribution of deaths events shown in Figure 4. The complete  $\mathcal{R}$  code is found in the Appendix. The path  $Y(n)$  is one realisation of the random processes and so several paths of the process are simulated and the mean path is taken.

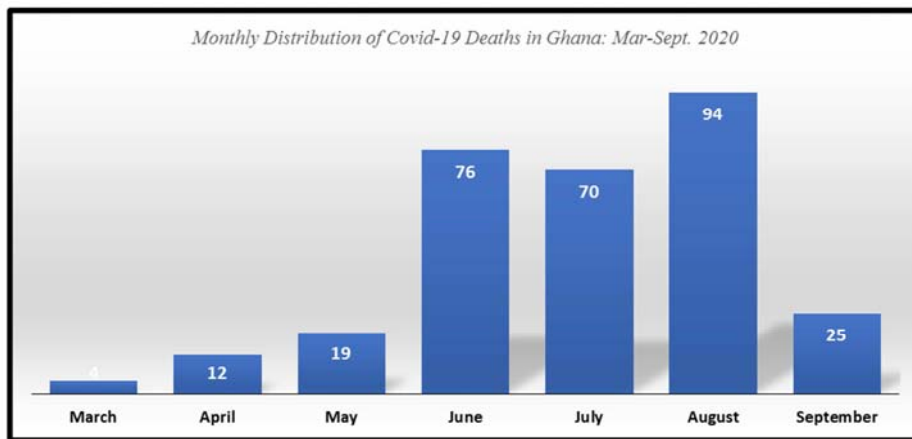


Figure 4. Monthly distribution of Covid-19 deaths in Ghana, March–September 2020.

## 4. Results and Discussion

Figure 5 shows the plots of the simulated and the actual daily death from Covid-19 in Ghana from March to August 2020. The cumulative number of deaths within the same period (March–August) is also shown in Figure 6. The simulated path using the proposed model predicts that there will be 268 deaths by August 31, 2020. The realised or the actual number of cumulative deaths at 31 August 2020 was 275. This shows that the model underpredict the number of deaths by 7 which translate as an error of 2.5%.

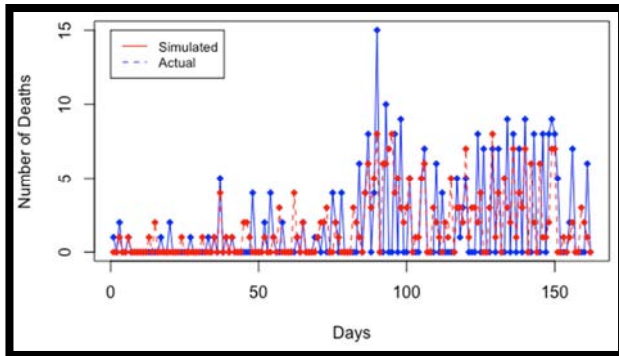


Figure 5. Actual vs. Simulated daily deaths in Ghana: March – August 2020.

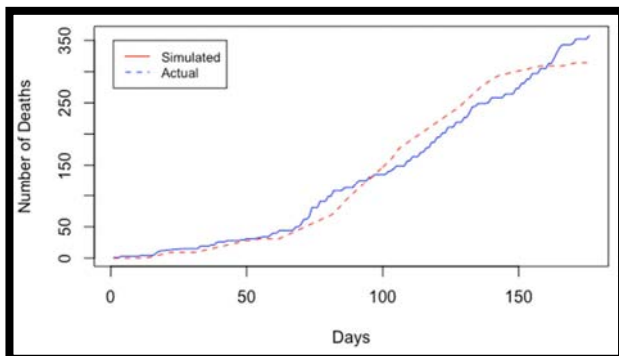


Figure 6. Actual vs. Simulated cumulative deaths in Ghana: March – August 2020.

The actual cumulative deaths and the simulated deaths, projected to February 2021 is shown in Figure 7. The cumulative deaths from the simulated and the actual daily deaths from Covid-19 from March to August 2020 is shown in Figure 8. The graph of cumulative simulated daily deaths is projected to the end of February 28, 2021 also in Figure 7. The results show that in the absence of any mitigating factors the number of Covid-19 deaths are expected to reach 446 by

February 28, 2021.

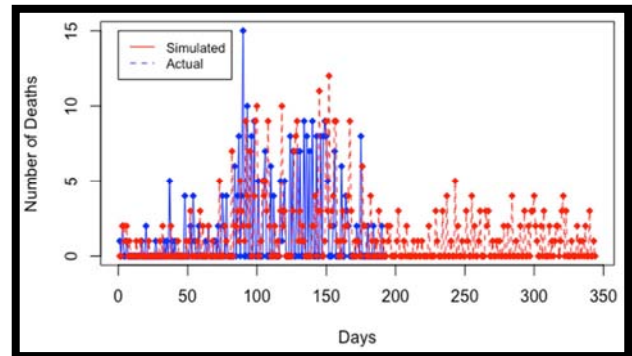


Figure 7. Projected Daily deaths from Covid-19: March 2020– February 2021.

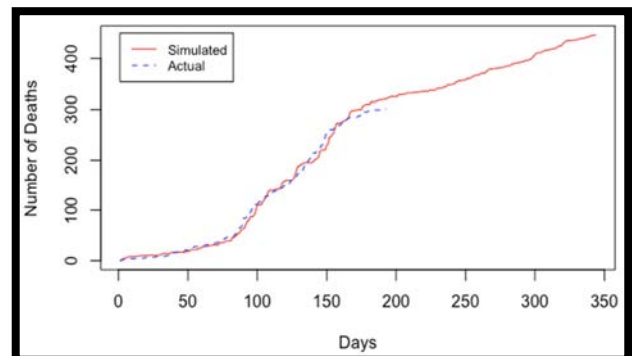


Figure 8. Projected cumulative deaths from covid-19: March – February 2021.

We emphasize here that although our primary interest in the model is its predictive capacity, we are also interested in the provision of a mathematical framework in the form of an explicit formula, which provides quantitative measures of the extent of the pandemic. To this end, we also obtained the summary statistics of actual and simulated number of deaths which is given in Table 1. The result shows that the simulated mean daily deaths March to August was 1.65. However, the actual mean number of daily deaths during this period was 1.69. In reality, since the number of deaths is integer valued it follows that both the actual and simulated can be approximated as 2 deaths per day. The  $t$ -test also shows that at 95% confidence interval, there is no significant difference between the means of actual and simulated number of deaths.

Table 1. Summary statistics of actual and simulated deaths.

	Actual Deaths	Simulated Deaths
Maximum event deaths	15	9
Minimum event deaths	1	1
Welch $t$ -test for difference between means		
$t$	$df$	95 % Confidence Interval for means
Mean	1.697531	1.654321
0.14991	294.85	−0.5240414 0.6104612
Alternative hypothesis: true difference in means is not equal to 0	$p$ -value = 0.8809	$p$ -value = 0.8809 > 0.05. We fail to reject the null hypothesis and conclude that, the means of the actual and Simulated are the same

## 5. Conclusion

In this paper, we have derived an explicit model based on the compound Poisson process, a class of the Levy processes to estimate fatalities from Covid-19 infections in Ghana. This is the first formal mathematical model in Ghana to analyse and make a short term forecast about the number of Covid-19 related deaths using stochastic processes. The model predicts the daily death toll as well as the total number of deaths in Ghana related to Covid-19. The results are obtained by fitting the simulated daily death cases as well as the cumulative number of deaths to the existing Covid-19 death data. The results show that the simulated cumulative deaths is in good agreement with the empirical data of actual deaths from Covid-19 and gives an accurate prediction of the future behaviour of deaths from pandemic in Ghana. We must however reiterate that while the short-term forecasts are in good agreement with the existing data, the current methodology is not able to estimate the asymptomatic behaviour that will eventually characterise the long-term distribution of the death process. To do this we need further information which is currently unavailable. In the future however, we believe we can obtain this information and make the necessary adjustment in the model to be able to make long-term predictions on the death situation.

## Appendix

*R*-Codes for Simulation

```
PPgen = function(lambda){
  X = 0
  Sum = 0
  flag = 0
  while (flag == 0){
    A = -log(runif(1))
    Sum = Sum + A
    if (Sum < lambda) { X = X + 1 } else { flag = 1 }
  } return(X) }
CPP = function(lambda,T,N) {
  h = T/N
  t = (0:T)/N
  X = rep(0,N + 1)
  F = rep(0,N + 1)
  I = rep(0,N)
  X[1] = 0
  for(i in 1:N) {
    I[i] = PPgen(h * lambda)
    if (I[i] == 0){F[i] = 0} else {F[i] = rnorm(1)}
    X[i + 1] = X[i] + F[i]}
  return(X) }
dr <- -CPP(5,71,299)
dr2 <- -round(dr,digits = 0)
dr1 <- -abs(dr2)
dr1
sum(dr1)
length(dr1)
```

## References

- [1] World Health Organization W.H.O. (accessed 12 August 2020). Coronavirus disease (COVID-2019) Situation reports. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports>
- [2] The Guardian Newspaper, <https://www.theguardian.com/world/2020/oct/11/coronavirus-world-map-which-countries-have-the-most-covid-cases-and-deaths> (accessed 12 August 2020)
- [3] Tom Orlik, Jamie Rush, Maeva Cousin, & Jinshan Hong, (2020). Corona virus Could Cost the Global Economy \$2.7 Trillion. Here's How Bloomberg. (2020). <https://www.bloomberg.com/graphics/2020-coronavirus-pandemic-global-economic-risk/>
- [4] Castorina P, Iorio A., & Lanteri D. (2020). Data analysis on coronavirus spreading by macroscopic growth laws. arXiv preprint arXiv:2003.00507.
- [5] Dehkordi A. H, Alizadeh M, Derakhshan P, Babazadeh P., & Jahandideh A. (2020). Understanding epidemic data and statistics: a case study of COVID-19. arXiv preprint arXiv:2003.06933.
- [6] Li R, Pei S, Chen B, Song Y, Zhang T, Yang W., & Shaman J. (2020). Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV2). *Science* 368(6490):489–493 DOI 10.1126/science. abb3221.
- [7] Elena Loli Piccolomini & Fabiana Zama (2020). Preliminary analysis of COVID-19 spread in Italy with an adaptive SEIRD model.
- [8] J. Kucharski, T. W. Russell, C. Diamond, Y. Liu, J. Edmunds, S. Funk, & R. M. Eggo. Early dynamics of transmission and control of covid-19: a mathematical modelling study.
- [9] D. Berger, K. Herkenhoff, & S. Mongey. (2020). An seir infectious disease model with testing and conditional quarantine. Technical report, Federal Reserve Bank of Minneapolis.
- [10] Walker P. G. T., Whittaker C., O. Watson, M. Baguelin, Ainslie K. E. C., S., Bhatia, S. Bhatt, A. Boonyasiri, O. Boyd, L. Cattarino, Z. Cucunuba, G. Cuomo-Dannenburg, A. Dighe, C. A. Donnelly, I. Dorigatti, S. van Elsland, R. FitzJohn, S. Flaxman, H. Fu, K. Gaythorpe, L. Gei-delberg, N. Grassly, W. Green, A. Hamlet, K. Hauck, D. Haw, S. Hayes, W. Hinsley, N. Imai, D. Jorgensen, E. Knock, D. Laydon, S. Mishra, G. Nedjati-Gilani, L. C. Okell, S. Riley, H. Thompson, J. Unwin, R. Verity, M. Vollmer, C. Walters, H. W. Wang, Y. Wang, P. Win-skill, X. Xi, N. M. Ferguson, & A. C. Ghani. (2020). The global impact of covid-19 and strategies for mitigation and suppression. Technical report, Imperial College.
- [11] Richards F.J. (1959). A flexible growth function for empirical use. *Journal of experimental Botany*, 10(2):290–301.
- [12] Hsieh Y-H. (2009). Richards model: a simple procedure for real-time prediction of outbreak severity. In: Ma Z, Zhou Y, eds. *Modeling and Dynamics of Infectious Diseases*. Singapore: World Scientific, 216–236.

- [13] Ivorra B, Ferrández M. R, Vela-Pérez M & Ramos A.M. (2019). Mathematical modeling of the spread of the coronavirus disease (COVID-19) taking into account the undetected infections. The case of China. *Commun Nonlinear Sci Numer Simul.* 2020 doi: 10.1016/j.cnsns.2020.105303.
- [14] Sameni R. (2020). Mathematical modeling of epidemic diseases; a case study of the COVID-19 coronavirus. arXiv preprint arXiv:2003.11371.
- [15] Crokidakis N. (2020). Data analysis and modelling of the evolution of COVID-19 in Brazil. arXiv preprint arXiv:2003.12150.
- [16] Saulo B. Bastos & Daniel O. Cajueiro. (2020). Modeling and forecasting the covid-19 pandemic in Brazil. arXiv preprint arXiv:2003.14288.
- [17] Hakan Andersson & Tom Britton. (2000). Statistical epidemic models and statistical analysis, Springer-Verlag, New York.
- [18] Roman Cherniha, Vasyl Davydovych Cherniha R. & Davydovych, V., (2020). A mathematical model for the COVID-19 outbreak. arXiv arXiv:2004.01487v2.
- [19] Hoang Pham Pham, H. (2020a). On Estimating the Number of Deaths Related to Covid-19. *Mathematics*, 8, 655.
- [20] Hoang Pham Pham. (2020b). Predictive Modelling of the Number of Covid-19 Death Toll in the United States Considering the Effects of Coronavirus-Related Changes and Covid-19 Recovered Cases, *International Journal of Mathematical, Engineering and Management Sciences* doi: 10.33889/IJMEMS.5.6.087.
- [21] M. Yousaf, S. Zahir, M. Riaz, Statistical analysis of forecasting COVID-19 for upcoming month in Pakistan, *Chaos Solitons Fractals* (2020), p. 109926, 10.1016/j.chaos.2020.109926.
- [22] Cont R., & Tankov P. (2009). Financial Modelling with jump processes, Chapman & Hall, New York.
- [23] Worldometer-covid-19 data. (Accessed: August, 2020). <https://www.worldometers.info/coronavirus/>.