Stochastic Modeling of an Infectious Disease
Part I: Understand the Negative Binomial Distribution
and Predict an Epidemic More Reliably

Hisashi Kobayashi[*]
Dept. of Electrical Engineering
Princeton University
Princeton, NJ 08544, U.S.A.

May 28, 2020

Summery

Why are the epidemic patterns of COVID-19 so different among different cities or countries which are similar in their populations, medical infrastructures, and people's behavior? Why are forecasts or predictions made by so-called experts often grossly wrong, concerning the numbers of people who get infected or die?

The purpose of this study is to better understand the stochastic nature of an epidemic disease such as COVID-19, and answer the above questions. The author hopes that this article will provoke discussions among the "modeling communities" and stimulate mathematically inclined people to study this interesting and important field, i.e., mathematical epidemiology.

Much of the work on infectious diseases has been based on "SIR deterministic models," pioneered by Kermack and McKendrick in 1927. In our study we will explore several stochastic models that can capture the essence of the seemingly erratic behavior of an infectious disease, which the deterministic model cannot explain. A stochastic model, in its formulation, takes into account the random nature of an infectious disease. Thus, such a model, if properly constructed, should be able to provide a more reliable and informative forecast of an epidemic pattern.

The stochastic model we study in this article is based on the ***birth-and-death process with immigration***(BDI for short), which was originally proposed in the study of population growth or extinction of some biological species. To the best of this author's knowledge, the BDI process model has not been investigated by the epidemiology community, perhaps for the reason we briefly discuss in Section 2.

The general birth-and-death (BD) process usually defies an attempt to obtain a closed solution for the time-dependent (i.e., transient) probability distribution of the population size, etc. The BDI process, however, is among a small number of BD processes, which we can solve analytically. An important feature of the BDI process is that its probability distribution function is a generalized

---

[*]The Sherman Fairchild University Professor of Electrical Engineering and Computer Science, Emeritus. Email: hisashi@princeton.edu, Blog: `http://hp.hisashikobayashi.com`, Wikipedia: `https://en.wikipedia.org/wiki/Hisashi_Kobayashi`

negative binomial distribution (NBD), with its parameter r being less than one. We show that the "coefficient of variation" (the standard deviation normalized by the mean) of the BDI process is larger than $r^{-1} > 1$. Furthermore, a NBD with small r has a long tail in its distribution form, like the zeta distribution (aka Zipf's law). These properties of the infection process explain why actual infection patterns exhibit enormously large variations. Furthermore, its mean value provided by a deterministic model is far from the median of the distribution. This explains why any forecast or prediction based on a deterministic model will fail more often than not.

In Part II of this report [1], we will present results of our extensive simulation study and further analysis of the stochastic model based on the BDI process.

Keywords: Infectious disease, COVID-19, Forecast and prediction, Stochastic model, Deterministic model, Kermack-McKendrick's SIR model, Basic and effective reproduction numbers, Birth-and-death process with immigration (BDI), Probability generating function (PGF), Partial differential equation (PDE), Negative binomial distribution (NBD), Coefficient of variation (CV), Compound Poisson process, Fisher's logarithmic distribution.

Contents

1   Introduction

Most of the mathematical models of infectious diseases seem to be based on the Kermack-McKendrick model published in 1927 [2], which was proposed to explain the rapid rise and fall in the number of infected population observed in epidemics such as the great plague in London where more than 15% of the population died (1665-66); and the cholera outbreak in London caused by contamination in the Thames River (1865), and the plague epidemic in Bombay (1906) [3]. The model consists of a system of three coupled nonlinear ordinary differential equations for the infected population $I(t)$, the susceptible population $S(t)$, and the recovered population $R(t)$. Kermack-McKendrick's **SIR model** is a **deterministic model**, which provides the **expected values** of these processes, which we denote as $\overline{S}(t)(= \mathsf{E}[S(t)]$, $\overline{I}(t)(= \mathsf{E}[I(t)]$ and $\overline{R}(t)(= \mathsf{E}[R(t)])$. A majority of biological and epidemiological models [4, 5, 6] fall in this class of deterministic models.

Actual observed data of the infected population, for instance, is merely an instance or **a sample path** of this stochastic process $I(t)$. The process naturally deviates from the expected value $\overline{I}(t)$ obtained by a deterministic model. Thus, a deterministic model alone fails to provide any quantitative explanation when observed data differ significantly from the expected value.

In a **stochastic** (or **probabilistic**) model, on the other hand, the intrinsic stochastic nature of a process is explicitly taken into account in its model formulation. The importance of stochastic processes in relation to problems of population growth was pointed out by W. Feller in 1939 [7]. He considered the **birth-and-death** process in which the expected birth and death rates (per person per unit time) were constants, say, λ and μ. D. G. Kendall [8] extended Feller's birth-and-death (BD) process by considering the birth and death rates as any specified functions of the time t, $\lambda(t)$ and $\mu(t)$. The BD process is a special class of *time-continuous discrete-state Markov process*, and has found applications in many scientific and engineering fields, including population biology [9], teletraffic and queueing theory [10], [11], [12], system modeling [13][14], pp. 63-94, [15], pp. 407-410.

I have done some investigation, with help from Prof. Hideaki Takagi, whose unpublished lecture note at the University of Tsukuba [16] provided me with several references, as to who coined the term "birth-and-death process with immigration," and have found the English statistician M.S. Bartlett (1910-2002) in his famous book "An Introduction to Stochastic Processes" [17] (1st edition in 1955) in his discussion of the birth-and-death process in "Section 3.4 Multiplicative chain: subsection 3.4.1 "The effect of immigration,"that he uses the phrase "a birth-death-and-immigration process." But Bartlett's doctoral student, David G. Kendall (1918-2007) gives a detailed analysis of the BDI process in his 1949 article [18], which Bartlett refers to in his 1949 article [19]. So my tentative conclusion was that Kendall was the first that worked on the BDI although he did not use the term "birth-death-immigration" or something to that effect.

Another English statistician, Norman T. J. Bailey published in 1964 "The Elements of Stochastic Process with Applications to the Natural Sciences," [20], and in "Section 8.7: The effect of immigration" (pp.97-101) he gives a thorough treatment of the BDI process. Linda J. S. Allen, "An Introduction to Stochastic Processes with Applications to Biology," (2nd Edition, 2011) [21] discusses the BDI process in Section 6.4.4: Simple Birth and Death with Immigration (pp. 254-258), but her focus seems to be more on the stable case. Frank P. Kelly [12] gives a brief treatment, providing the steady-state distribution. All other numerous textbooks on random processes make no mention of the BDI process, and even the above handful of authors who might have had epidemiologists in mind among their readership seem to treat the BDI process for its possible application to population biology, and none allude to its use in epidemiology.

2 A Brief Review of SIR Deterministic Model

As stated earlier, a majority of mathematical models reported in the literature on infectious diseases have been deterministic models, following the pioneering work of almost a century ago by W.G. Kermack and A.G. McKendrick [2] [1]. They assumed

1. At any time t, an individual is either susceptible (S), infected and infectious (I) or recovered and immune (R).

2. Only susceptible individuals can get infected, remain infectious for some time, and recover and become completely immune.

3. There are no births, deaths, immigration or emigration during the study period. In other words, the community is *closed*.

Consequently, individuals can only make two types of transitions: (i) from S to I, and (ii) from I to R. Thus, the Kermack-Mckendrick model is often referred to as an *SIR* epidemic model. A model which assumes no immunity (i.e., a recovered person becomes immediately susceptible) is called an *SIS* model. If we explicitly consider an *exposed* state, during which an infected individual is not yet infectious, the model is called an *SEIR* model. A model in which immunity wanes after some period is called an *SIRS* model, and so forth. In the remainder of this section we give a brief account of the SIR model in a closed community so that the reader can compare this deterministic model to our stochastic model to be presented in the next section. For details of the Kermack-McKedrick type deterministic models, the readers are referred to abundant books and articles; Anderson and May, [5], Martcheva [6] to name just a few.

Let $\overline{S}(t)$, $\overline{I}(t)$ and $\overline{R}(t)$, respectively[2], denote the number of the susceptible, infected and recovered at time t. Since we assume no births, deaths, immigration nor emigration, we have

$$\overline{S}(t) + \overline{I}(t) + \overline{R}(t) = N, \quad \text{for all} \ \ t \geq 0, \tag{1}$$

where N is a constant number, representing the population of the community.

From the set of assumptions stated above, the deterministic processes can be defined by the following set of three differential equations:

$$\frac{d\overline{S}(t)}{dt} = -\beta \overline{S}(t)\overline{I}(t), \tag{2}$$

$$\frac{d\overline{I}(t)}{dt} = \beta \overline{S}(t)\overline{I}(t) - \mu \overline{I}(t), \tag{3}$$

$$\frac{d\overline{R}(t)}{dt} = \mu \overline{I}(t). \tag{4}$$

These differential equations, together with (1) and the initial condition

$$\overline{I}(0) = I_0, \quad \text{and} \ \ \overline{R}(0) = 0, \tag{5}$$

---

[1]William Ogilvy Kermack (1898-1970) was a Scottish biochemist and Anderson Gray McKendrick (1876-1943) was a Scottish military physician and epidemiologist.

[2]We adopt this notation to distinguish them from $S(t)$, $I(t)$ and $R(t)$ which are stochastic processes, as used in other parts of this article. The time functions used in deterministic models usually correspond to the expectation or stochastic mean of the corresponding stochastic processes.

define the deterministic model. It is easy to see that $\overline{S}(t)$ is monotone decreasing, and $\overline{R}(t)$ is monotone increasing. The function $\overline{I}(t)$ increases or decreases at time t, depending on whether the ratio R_t defined by the following expression is greater or smaller than unity.

$$\mathcal{R}_t = \frac{\beta \overline{S}(t)}{\gamma} \leq \frac{\beta \overline{S}(0)}{\gamma} = \mathcal{R}_0 \tag{6}$$

Its initial value $\mathcal{R}_0$ is referred to as the ***basic reproduction number***, a term having its origin in demography. The ratio R_t is called the ***effective reproduction number.*** or the ***real-time reproduction number*** and is more meaningful than $\mathcal{R}_0$, which is a static number, in estimating the current epidemic situation and making a decision to control the epidemic.

The term $\beta \overline{S}(t)\overline{I}(t)$ in (2) and (3) comes from the argument that the susceptible must have contact with the infected in order to get infected, and if we assume some sort of *uniform mixing*, the infections should occur at a rate proportional to $\overline{S}(t)\overline{I}(t)$. Consequently, the unit of the parameter β is [person/unit-time/person·person], whereas the other parameter [3] μ has the unit of [person/unit-time/person].

When the total population N of (1) is sufficiently large, $S(t) \approx N$ is much larger than $I(t)$ and can be treated as unchanged, at least in the initial phase of an epidemic. Then by setting

$$\lambda \approx \beta S(t) \approx \beta S(0) \; [\text{person/unit-time} \cdot \text{person}], \tag{7}$$

and by substituting this into (3), we have the following ordinary differential equation (ODE):

$$\frac{d\overline{I}(t)}{dt} = (\lambda - \mu)\overline{I}(t), \tag{8}$$

from which and the initial condition (5), we readily find the solution for $\overline{I}(t)$:

$$\overline{I}(t) = I_0 e^{at}, \;\; t \geq 0, \;\; \text{where} \;\; a = \lambda - \mu. \tag{9}$$

which is an exponentially growing or decaying function, depending on whether $a > 0$ or $a < 0$. When $a = 0$, it is a constant I_0 for all $t \geq 0$.

When the infected $I(t)$ grows to the extent that the approximation (7) no longer holds, i.e., the "infinite population" assumption fails, we have to deal with the nonlinear differential equations of (2) and (3). A major advantage of the SIR model is that because of the product term $S(t)I(t)$, the differential equations take into account explicitly the fact that occurrences of infections will gradually decreases towards to zero as the susceptible population becomes extinct towards the end of the infection processes. The main drawback of the SIR model, on the other hand, is its inability to capture any probabilistic fluctuation of the infection process. The SIR model may be an appropriate model in describing the interactions between the susceptible group and infected group in a closed environment, such as a hospital, retirement home, cruise ship, night club, etc. But it is a poor model in describing a major outbreak of an epidemic in a larger environment such as a city or a country, where most infections take place independently and randomly, and the product term $\beta \overline{S}(t)\overline{I}(t)$ does not have any significant meaning. Furthermore, the product term makes the entire SIR model a nonlinear system, and makes the system mathematically intractable, excepting for a few simple cases, which may not be useful in reality.

---

[3] Another Greek letter γ is often used for the recovery rate, but we use μ to be consistent with our notation in the next section.

As we will show in this article and Part II [1], the deterministic model, in addition to being unable to describe the stochastic fluctuation of an epidemic pattern, is more likely to grossly overestimate the number of casualties. Thus, the deterministic model is not only limited in its applicability, but can be damaging and harmful in some cases.

3   A Stochastic Model for an Infectious Disease

In this and following sections we will discuss our stochastic model based on the **birth-and-beath with immigration** (BDI) process. It is a special case of general birth-and-death (BD) process[4]. Before we present a detailed mathematical analysis of this model, we will show an example of our simulation model based on the BDI process.   Figure 1 shows the first 6 out of a total of 12 simulation
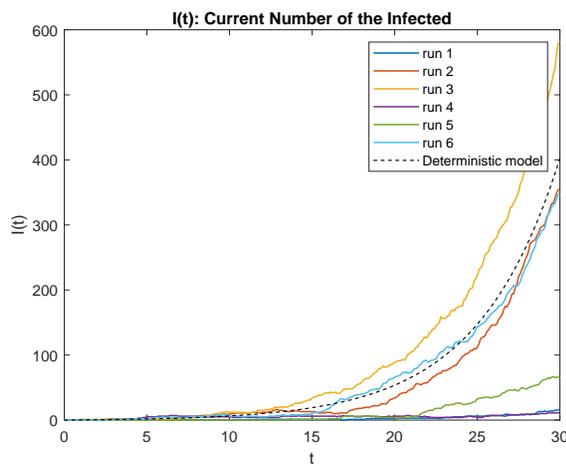

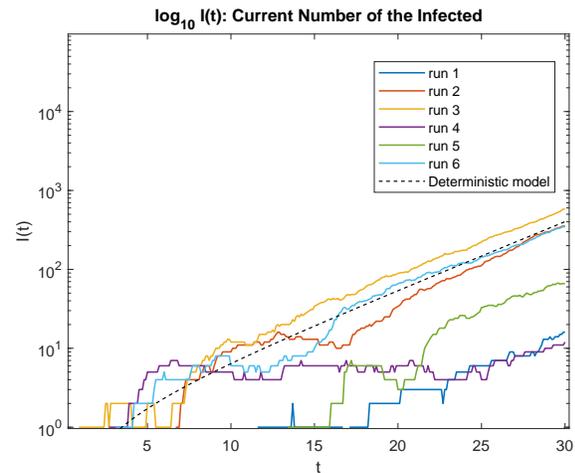Figure 1: The first 6 simulation runs of BDI process and a deterministic model curve.

Figure 2: Semi-log plots of the same runs. The baseline of y axis is $10^0 = 1$.

runs consecutively done in one execution of our simulation program in a MATLAB script. Our simulator is based on the *event scheduling approach*[5], where events are arrivals of infected persons from the outside (at rate ν persons/day), occurrences of secondary infections within the community (at rate λ infections/day/infectious person), and recovery/deaths of the infected persons (at rate μ recovery/death/day/infected person). The parameter values of $\lambda = 0.3, \mu = 0.1$ and $\nu = 0.2$ are assumed. Figure 2 plots the same set of curves in a semilog scale. The exponential growth curves are shown as straight-lines; the initial part of the simulation is more clearly shown in the semilog scale. Figures 3 and 4 are the plots of the remaining six runs.

There are at least two questions concerning these simulation runs.

1. Why are the variations among different simulation runs so large?

---

[4]William (Willy) Feller introduced what is now known as the "birth-and-death" process in his 1939 article [7] published in German regarding the problem of population growth. He used the term "Vermehrung" (reproduction) and "Tod" (death). David. G. Kendall in his 1948 paper [8] referred to Feller's model as "birth-and-death" process. Feller used this term in his Volume I [9], whose first edition was in 1950.

[5]The time-asynchronous event scheduling approach is a more time efficient and accurate simulation method than a time-synchronous approach. See e.g., [13] pp. 230-234, or [14], pp. 626-630.
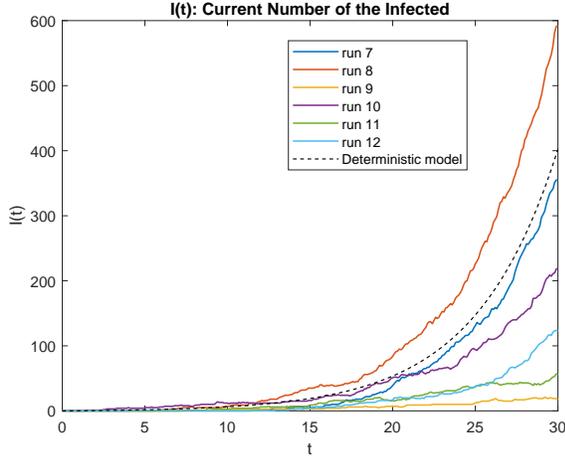
Figure 3: The next 6 simulation runs of BDI process and a deterministic model curve.
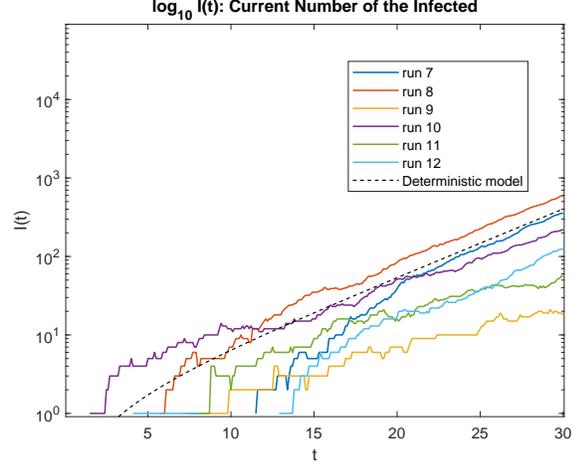
Figure 4: Semi-log plots of the same runs. The baseline of y axis is $10^0 = 1$.

2. There are more runs whose plots are below the deterministic model curve. Does the deterministic model tend to overestimate the size of the infected population? If so, why?

The analysis of the BDI process in the following sections should be able to answer these questions.

3.1 Formulation for the time-dependent solution for the stochastic model

Let $I(t)$ represent the number of infected persons at time t, and $P_n(t)$ be the time-dependent (or transient) probability mass function (PMF)[6] of the process $I(t)$, i.e.,

$$P_n(t) = \Pr[I(t) = n], \quad n = 0, 1, 2, \cdots, \quad \text{and} \quad t \geq 0. \tag{10}$$

We assume that each infected person is infectious, and infects susceptible persons at rate λ [persons/unit-time/person], where the time unit can be arbitrary, e.g., a second, an hour, a day, etc. Let us assume that an infected person recovers, gets removed or dies at rate μ/[unit-time]. Thus, μ^{-1} [unit-time] is the *mean infectious period*. The ratio λ/μ is equal to the *basic reproduction number*, i.e., the mean number of infections caused by an infected person during the infectious period.

We can formulate an infectious disease as a birth-and-death (BD) process, by defining the birth and death rates both of which are simple linear functions of the state n of the process $I(t)$:

$$\begin{aligned}
\lambda_n &= n\lambda + \nu, \\
\mu_n &= n\mu,
\end{aligned} \tag{11}$$

A few remarks are in order. This particular state-dependent BD process is also known as the "birth-death-immigration (BDI)" process (see e.g.,[12], p. 14 for the steady state distribution), in which

---

[6]We could use perhaps a more common term the *probability distribution function* but PMF is more explicit that we are dealing with a discrete distribution, not a continuous distribution.

7

the **parameters** λ, μ **and** ν represent the *birth* (i.e., secondary infection), *death* (i.e., recovery or death) and *immigration* (i.e., arrival of an infected individual from outside) *rates*, respectively.

A few remarks are in order.

1. In actuality, "recovery," "removal" and "death" are distinctly different matters. In analyzing the infection process, however, these three sources of loss from the susceptible or infected population, are mathematically equivalent in the sense they will not contribute to the infection process in the future. We assume here that those who have recovered from the disease have acquired immunity and will not be susceptible nor infectious.

    The assumption that each infected individual recovers (or is removed or dies) at rate μ is equivalent to assuming that the duration S that each sick person remains infectious is exponentially distributed with mean $1/\mu$, that is:

$$\Pr[S \le s] = 1 - e^{-\mu s}, \quad s \ge 0. \tag{12}$$

2. It can be shown mathematically that many of our results to be obtained in terms of the probability mass function (PMF) of $I(t)$, and other related quantities are *insensitive* to the actual distribution of S . All that matters is that we set $\mu = \overline{S}^{-1}$, where $\overline{S} = \mathsf{E}[S]$.

3. In the present paper, we assume that the population is homogeneous, and the susceptible population size remains sufficiently large, thus mathematically treated as "infinite." Furthermore, the parameters λ, μ and ν are assumed to be constant. Many of our results can be extended to the case of multiple types of populations (e.g., clustering of infections): a model with the susceptible population decreases as some of them get infected; and the case where the model parameters' values change (e.g., the situation where the infectious rate λ may change as people's behavior changes), and these generalized models will be discussed in subsequent reports.

We can show (see e.g.,[15], pp.407-408) that the PMF (10) of this BD process should satisfy the following set of linear differential-difference equations, a.k.a. *Kolmogorov's forward equation*:

$$\frac{dP_0(t)}{dt} = -\nu P_0(t) + \mu P_1(t)$$

$$\frac{dP_n(t)}{dt} = ((n-1)\lambda + \nu)\, P_{n-1}(t) - (n(\lambda + \mu) + \nu)\, P_n(t) + (n+1)\mu P_{n+1}(t), \quad n = 1, 2, 3 \cdots \tag{13}$$

with the initial condition

$$I(0) = I_0, \text{ i.e., } P_n(0) = \delta_{n,I_0}, \quad n = 0, 1, 2, \cdots, \tag{14}$$

where $\delta_{m,n}$ is Kronecker's delta.

We transform the above set of infinitely many equations (13) into a single equation by using the *probability generating function* (PGF) (see e.g.,[15], p. 402) defined by

$$G(z,t) = \mathsf{E}[z^{I(t)}] = \sum_{n=0}^{\infty} z^n P_n(t). \tag{15}$$

Multiply the set of equations (13) by z^n and sum them from $n = 0$ to ∞, obtaining the following partial differential equation:

$$\frac{\partial G(z,t)}{\partial t} = (z-1)\left[(\lambda z - \mu)\frac{\partial G(z,t)}{\partial z} + \nu G(z,t)\right], \tag{16}$$

with the boundary condition

$$G(z,0) = z^{I_0}. \tag{17}$$

3.2 Stochastic means of the infected process *I(t)* and related processes

Although the process $I(t)$ is the main focus of our analysis, it will be worthwhile to introduce related processes and our assumptions.

Definition 1.

1. *The process $A(t)$ is the cumulative count of external arrivals of infectious individuals from the outside. We assume that $A(t)$ is a Poisson process with rate ν [persons/unit time],*

2. *The process $B(t)$ is the cumulative count of internally infected individuals. We assume that the birth of such persons occurs at the rate of λ [persons/unit time/infectious person].*

3. *The process $R(t)$ is the cumulative count of recovered/removed or dead individuals. We assume that the departure of such persons occurs at the rate of μ [persons/unit time/infected person]. Note that all infected persons are infectious persons until their recovery/removal/death.*

4. *The process $I(t)$ is the present number of infected persons, i.e.,*

$$I(t) = I(0) + A(t) + B(t) - R(t). \tag{18}$$

□

The expectation and variance of the above processes will be of our interest, which we denote by

$$\overline{A}(t) = \mathsf{E}[A(t)], \quad \overline{B}(t) = \mathsf{E}[B(t)], \quad \overline{R}(t) = \mathsf{E}[R(t)], \quad \text{and} \quad \overline{I}(t) = \mathsf{E}[I(t)], \tag{19}$$
$$\sigma_A^2(t) = E[(A(t) - \overline{A}(t))^2], \quad \text{etc.} \tag{20}$$

Before we discuss how to find the PGF $G(z,t)$ from the PDE (16), let us derive first an ordinary differential equation for $\overline{I}(t)$. By dividing both sides of (16) by $(z-1)G(z,t)$, we will have

$$(z-1)^{-1}\frac{\partial \ln G(z,t)}{\partial t} = (\lambda z - \mu)\frac{\partial \ln G(z,t)}{\partial z} + \nu. \tag{21}$$

By setting $z = 1$, we find[7]

$$\frac{d\overline{I}(t)}{dt} = (\lambda - \mu)\overline{I}(t) + \nu. \tag{22}$$

---

[7]Alternatively, we can obtain this differential equation directly, by multiplying each equation in (13) by n and summing them up from $n = 0$ to infinity.
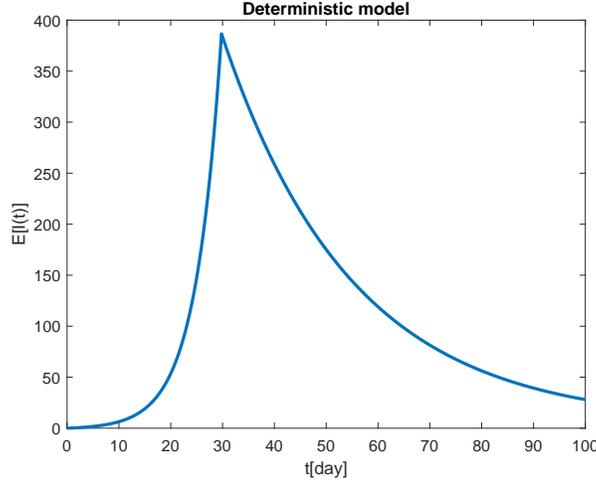
Figure 5: $\overline{I}(t)$ when $\lambda = 0.3, \mu = 0.1, \nu = 0.2$ for $0 \le t \le t_1 = 30$ at which point λ is changed to $\lambda' = 0.06$.

where, on the LHS[8], we first set $z = 1$ (which corresponds to differentiation at $z = 1$), and use L'Hôpital's rule, obtaining [9] $\mathsf{E}[I(t)] = \overline{I}(t)$. The ordinary differential equation (22) can be solved, yielding

$$\overline{I}(t) = I_0 e^{at} + \frac{\nu}{a}\left(e^{at} - 1\right), \quad t \ge 0, \quad \text{where} \quad I_0 = I(0) \quad \text{and} \quad a = \lambda - \mu. \tag{23}$$

If the model parameters are set to new values, say, to λ', μ' and ν' at some point $t = t_1 \ge 0$, then the solution $\overline{I}(t)$ for $t \ge t_1$ is given by

$$\overline{I}(t) = I_1 e^{a'(t-t_1)} + \frac{\nu'}{a'}\left(e^{a'(t-t_1)} - 1\right), \quad t \ge t_1, \quad \text{where} \quad I_1 = \overline{I}(t_1), \quad a' = \lambda' - \mu'. \tag{24}$$

It should be clear that $\overline{I}(t)$ diverges to infinity, if $a > 0$ in (23) and converges to $\nu/|a|$ in the limit $t \to \infty$ if $a < 0$. Similarly, in (24), the process converges to $\nu'/|a'|$, if $a' < 0$.

If $a = 0$, then[10]

$$\overline{I}(t) = I_0 + \nu t = I_0 + A(t) \ t \ge 0, \quad \text{when} \quad a = 0. \tag{25}$$

In Figure 5 we show the case where $\lambda = 0.3, \mu = 0.1$ and $\nu = 0.2$ and at $t_1 = 30$, a new parameter $\lambda' = 0.06$ is set, whereas the original values of μ and ν are retained.

The mean values of other processes can be easily found. $A(t)$ is a Poisson process with rate ν, which implies

$$\overline{A}(t) = \nu t, \quad \text{for all} \quad t \ge 0. \tag{26}$$

---

[8]The abbreviations LHS and RHS mean the left-hand side and right-hand side, respectively.

[9]Here we use an important property of PGF, i.e, $\frac{\partial G(z,t)}{\partial t} = \mathsf{E}[I(t)z^{I(t)-1}]$, and by setting $z = 1$, the RHS becomes $\mathsf{E}[I(t)]$. We changed the order of differentiation w.r.r. to z and t, which can be justified because the function $G(z,t)$ is an analytic function, i.e., it is continuous and differentiable everywhere.

[10]The second term becomes $0/0$, so we apply L'Hôpital's rule.

Since each person in the infected population $I(t)$ infects others at the rate of λ persons/unit time, the differential of $\overline{B}(t)$ is given by

$$\frac{d\overline{B}(t)}{dt} = \lambda \overline{I}(t). \tag{27}$$

From this and (23) we obtain

$$\overline{B}(t) = \lambda \int_0^t \overline{I}(u)\, du = \frac{\lambda I_0(e^{at} - 1)}{a} + \frac{\lambda \nu(e^{at} - 1)}{a^2} - \frac{\lambda \nu t}{a}$$
$$= \frac{\lambda}{a}(\overline{I}(t) - I_0 - \overline{A}(t)), \quad 0 \le t \le t_1. \tag{28}$$

For $t \ge t_1$, we find

$$B(t) = B(t_1) + \frac{\lambda' I_1(e^{a'(t-t_1)} - 1)}{a'} - \frac{\lambda' \nu'(t - t_1)}{a} + \frac{\lambda' \nu'(e^{a(t-t_1)} - 1)}{a^2}$$
$$= \frac{\lambda'}{a'}\left(I_1 e^{a'(t-t_1)} + \frac{\nu'}{a'}(e^{a'(t-t_1)} - 1) - I_0 - \nu' t\right)$$
$$= \frac{\lambda'}{a'}(I(t) - I_0 - \overline{A}'(t)), \quad \text{where } \overline{A}'(t) = \nu' t, \tag{29}$$

which takes the same form as that for $t \le t_1$. It should be worthwhile to note that both (28) and (29) could be derived from the mean value of the identity equation (18) together with the relation $\overline{R}(t) = \frac{\mu}{\lambda}\overline{B}(t)$, which is evident as shown below.

The recovery process $\overline{R}(t)$ should satisfy the following differential equation, similar to (27):

$$\frac{d\overline{R}(t)}{dt} = \mu \overline{I}(t). \tag{30}$$

Thus, it readily follows:

$$\overline{R}(t) = \begin{cases} \mu(\overline{I}(t) - I_0 - \overline{A}(t))/a, & 0 \le t \le t_1 \\ \mu'(\overline{I}(t) - I_0 - \overline{A}'(t))/a', & t \ge t_1. \end{cases} \tag{31}$$

The above expressions for $\overline{I}(t)$ and other processes can be viewed as our **deterministic (or non-probabilistic) model** for the dynamics of the BDI process. From these simple expressions, we can extract a few important characteristics concerning the mean values of $I(t)$ and others.

1. $a = \lambda - \mu$ determines the exponential growth or decay rate of $\overline{B}(t), \overline{R}(t)$ as well as $\overline{I}(t)$.

2. ν is merely a linear scaling factor for $\overline{I}(t)$ and other processes, and so is I_0, the initial number of the infected.[11]

3. If $a > 0$, then $\overline{I}(t)$ grows exponentially without bound; if $a < 0$, it decays exponentially towards $\nu/|a|$. If $a = 0$, $\overline{I}(t) = I_0 + \nu t$, i.e., $\overline{I}(t)$ grows linearly.

---

[11]Eq.(23) can be rewritten as $\overline{I}(t) = \left(I_0 + \frac{\nu}{a}\right)e^{at} - \frac{\nu}{a}$. So $\frac{\nu}{a} + I_0$ is the multiplying coefficient of the exponential term e^{at}.

4. The ratio of the infection rate (or reproduction rate) λ to the recovery or removal rate μ

$$R_0 = \frac{\lambda}{\mu} \tag{32}$$

is called the **basic reproduction number** in epidemiology (see e.g., [6], p. 21). The term was originally defined in the context of a deterministic model called the **SIS**(*Susceptible-Infected-Susceptible*) epidemic model. It is the average number of persons whom an infectious person infects before his/her recovery, removal, or death. The reproduction number determines whether the infection will grow exponentially, die out, or remain constant, depending on whether $R_0 > 1$, $R_0 < 1$, or $R_0 = 1$, respectively. The exponential parameter a can be expressed in terms of R_0 and μ:

$$a = \lambda - \mu = (R_0 - 1)\mu. \tag{33}$$

5. The amount of time T that takes for $\overline{I}(t)$ or other related quantities to double, and thes exponential growth parameter a are related by

$$e^{aT} = 2, \quad \text{or equivalently} \quad aT \approx 0.693. \tag{34}$$

Note that both the integral and derivatives of the exponential function e^{at} are also $\propto e^{at}$. Thus, the above formula for T equally applies, when cumulative numbers or incremental numbers are to be counted for a given a via an observed T.

6. Unless we can expect to increase the value of μ by improving the medical service or producing an effective vaccine to immunize the susceptible population, the only options we have for controlling an infectious disease is to **increase** μ by removing as many infectious individuals away from susceptible population as possible, and/or to **decrease** λ by increasing the so-called **social distances** between the susceptible and the infectious. We will provide an illustrative example in Part II [1].

Example 1 Consider the following community: The external arrival rate of infected individuals $\nu = 0.2$ [persons/day], i.e. one such such incidence every 5 days on average. The average number of days required for an infected person to recover, be removed or die, is $\mu^{-1} = 10$ [days], i.e., the recovery rate $\mu = 0.1$ [per day] for each infected person. Suppose that the average number of secondary infections caused by an infected individual is estimated as $R_0 = \lambda/\mu = 3$ for each infectious person. Then this value and μ provide an estimate of the infectious rate $\lambda = 0.3$ [persons/day/infected person]. Thus, the exponential growth parameter is estimated as $a = \lambda - \mu = 0.2 > 0$. One can check the validity of the model and/or consistency among the three estimated parameters, by computing a second estimate of the a from the formula (34). In Part II, we will discuss various ways of estimating the model parameters.

Once we have obtained reliable model parameters, we can predict the expected value of the infected process $\overline{I}(t)$, assuming that there was no infected person at $t = 0$, i.e., $I(0) = I_0 = 0$:

$$\overline{I}(t) = \frac{\nu}{a}\left[e^{at} - 1\right] = e^{0.2t} - 1. \tag{35}$$

Thus, this curve predicts the expected number of the infected population, excluding those who have recovered, been removed or have died.

$$\overline{I}(0) = 0, \quad \overline{I}(5) = 1.7, \quad \overline{I}(10) = 15.0, \quad \overline{I}(15) = 19.1, \quad \overline{I}(20) = 53.6, \quad \overline{I}(25) = 147.4,$$
$$\overline{I}(30) = 402.5, \quad \overline{I}(35) = 1,095.5, \quad \overline{I}(40) = 2,980.0, \quad \overline{I}(45) = 8,102.1, \quad \overline{I}(50) = 22,025.5.$$

$\square$

3.3 Steady-state distribution of the *I(t)*

So far we have discussed only the mean values of the random process $I(t)$ and other processes. Before we find the probability mass functions $P_n(t), n = 0, 1, 2, \cdots$ for any t, we obtain in this section the steady-state distribution $\lim_{t\to\infty} P_n(t) = \pi_n$, if it exists. We know already that when $a > 0$, such distribution does not exist.

Thus, the steady state distribution can possibly exist, only when $a \leq 0$. In order to find it, we set the LHS of the PDE (16) equal to zero, obtaining the following ordinary differential equation for the PGF $G(z, \infty)$:

$$(\lambda z - \mu)\frac{dG(z, \infty)}{dz} + \nu G(z, \infty) = 0, \tag{36}$$

which readily leads to

$$\frac{d \ln G(z, \infty)}{dz} = -\frac{\nu}{\lambda z - \mu}. \tag{37}$$

Integrating the above and using the boundary condition $G(1, t) = 1$ for any t, we find

$$G(z, \infty) = \left(\frac{1 - \frac{\lambda}{\mu}}{1 - \frac{\lambda}{\mu}z}\right)^r, \quad \text{where} \quad r = \frac{\nu}{\lambda}. \tag{38}$$

This PGF reminds us of the **negative binomial distribution** (NBD). This distribution was originally introduced to express the probability of the *number of failures n* needed to achieve r successes in a sequnce of Bernoulli trials, when the probability of failure is q.

Definition 2 (Negative binomial distribution (NBD)). *Negative binomial distribution NB(r, q) is defined by*

$$P_n^{NB} = \binom{n + r - 1}{n}(1 - q)^r q^n = \frac{\Gamma(n + r)}{n!\Gamma(r)}(1 - q)^r q^n, \quad n = 0, 1, 2, \cdots, \tag{39}$$

where the parameter r is a positive real number. $\square$

When r is a positive integer, and $0 < q < 1$ the above reduces to the classical definition of the (shifted)[12] negative binomial distribution, sometimes called the Pascal distribution, associated with Bernoulli trials. The Gamma function $\Gamma(x)$ is defined for a positive real number x by[13]

$$\Gamma(x) = \int_0^\infty y^{x-1}e^{-y}\,dy, \quad x > 0. \tag{40}$$

The *probability generating function* of NB(r, q) of (39) is given by

$$G(z) = \sum_{n=0}^\infty \binom{n + r - 1}{n}(1 - q)^r q^n z^n = \frac{1 - q}{1 - qz}, \quad |z| < q^{-1}. \tag{41}$$

The mean and variance of a RV (random variable) X possessing this distribution can be readily found:

$$\mathsf{E}[X] = \frac{qr}{1 - q}, \quad \text{and} \quad \text{Var}[X] = \frac{qr}{(1 - q)^2}. \tag{42}$$

From (38) and the above formula, we readily find the steady state distribution of $I(t)$ for $a < 0$ is given as follows [12],p.14.

$$\pi_n = \lim_{t\to\infty} P_n(t) = \binom{n + r - 1}{n}\left(\frac{\lambda}{\mu}\right)^n\left(1 - \frac{\lambda}{\mu}\right)^r. \tag{43}$$

---

[12]Some authors define the negative binomial distribution as the distribution of the *number of trials*, instead of the *number of failures*, needed to achieve r successes. Under this definition, $n = r, r + 1, r + 2, \cdots$. The probability distribution (39) is then referred to as the ***shifted negative binomial distribution*** (see e.g.,[15], pp. 59-62).

[13]This definition can be extended for a complex number z, with $\Re(z) > 0$.

4 Time-Dependent Probability Distribution of the Infected Process $I(t)$

The partial difference equation (PDE) (16) for the PGF $G(z,t)$ is a linear PDE and is sometimes referred to as **planar differential equation** (see [22], [14], pp. 600-605). This type of PDE can be solved by using **Lagrange's method** with its *auxiliary differential equations*, which is discussed in Appendix A. [14]

4.1 When the system is initially empty, i.e., $I(0) = 0$

When the system is initially empty, i.e., $I(0) = 0$, its PGF can be found by solving the PDE given by (see (A.20) of Appendix A)

$$G(z,t) = \left(\frac{a}{\lambda e^{at} - \mu - \lambda(e^{at} - \mu)z} \right)^r, \quad \text{where } a = \lambda - \mu, \text{ and } r = \frac{\nu}{\lambda}. \tag{44}$$

If we define a function $\beta(t)$

$$\beta(t) = \frac{\lambda(e^{at} - 1)}{\lambda e^{at} - \mu}, \tag{45}$$

we can write (44) compactly as

$$G(z,t) = \left(\frac{1 - \beta(t)}{1 - \beta(t)z} \right)^r, \tag{46}$$

which is the PGF of the generalized negative binomial distribution $\text{NB}(r, \beta(t))$, defined in Definition 2. Thus, we find that the PMF of the BDI process $I(t)$ is given by

$$P_n(t) = \Pr[I(t) = n] = \binom{n + r - 1}{n} (1 - \beta(t))^r \beta(t)^n, \quad n = 0, 1, 2, \cdots. \tag{47}$$

When $a = 0$, the above expression can be simplified. Noting

$$\lim_{a \to 0} \frac{e^{at} - 1}{a} = t, \tag{48}$$

we find

$$\beta(t) = \frac{\mu t}{\mu t + 1}, \quad \text{and} \quad 1 - \beta(t) = \frac{1}{\mu t + 1}, \tag{49}$$

which lead to

$$G(z,t) = \left(\frac{1}{1 + \mu t - \mu t z} \right)^r, \tag{50}$$

and

$$P_n(t) = \binom{n + r - 1}{n} \left(\frac{1}{\mu t + 1} \right)^r \left(\frac{\mu t}{\mu t + 1} \right)^n. \tag{51}$$

---

[14]It will be worth noting that Ren and Kobayashi [23], [24] discuss this type of PDE in the analysis of multiple on-off sources in traffic characterization of a data network.

The mean and variance can be computed from the distribution function obtained above. But we already know the expression for the mean $\bar{I}_{BDI}(t)$ from (23), which can be also found from the formula of the NB(r, q, i.e.,

$$\mathsf{E}[I_{BDI}(t)] = \frac{r\beta(t)}{1 - \beta(t)} = \frac{\nu}{\alpha}(e^{at} - 1), \tag{52}$$

which certainly agrees with (23). Similarly, the variance can be found from the formula of NB(r, q) as

$$\sigma^2_{BDI}(t) = \frac{r\beta(t)}{(1 - \beta(t))^2} = \frac{\nu(\lambda e^{at} - \mu)(e^{at} - 1)}{a^2}. \tag{53}$$

The reason why this simple innocent-looking equation is put in a box is because this expression for **the variance is perhaps the most important feature** behind the erratic behavior we observe in the COVID-19, or an infectious disease in general, as we analyze below and demonstrate by presenting simulation results in Part II.

4.2   The process $I(t)$ with an arbitrary initial condition

Up to now, we have focused on the case when the system is initially empty. In this section, we wish to study a general case where $I(0) = I_0$ is an arbitrary non-negative integer. The PGF solution for this general case is, in fact, given in (A.19) of Appendix A. In this section we will add to the PGF and PMF of this general case a subscript or superscript $_{BDI:I_0}$ wherever we need to distinguish similar functions of other processes. So we can write

$$G_{BDI:I_0}(z, t) = G_{BDI:I_0=0}(z, t) \ G_{BD:I_0}(z, t), \tag{54}$$

where

$$G_{BDI:I_0=0}(z, t) = \left(\frac{a}{\lambda e^{at} - \mu - \lambda(e^{at} - 1)z} \right)^r \tag{55}$$

$$G_{BD:I_0}(z, t) = \left(\frac{\mu e^{at} - \mu + (\lambda - \mu e^{at})z}{\lambda e^{at} - \mu - \lambda(e^{at} - 1)z} \right)^{I_0}, \tag{56}$$

where the first given by (55) is the same as (44) discussed in a preceding section, whereas the second one (56) is the PGF of **the birth-and-death** process, obtained by setting $\nu = 0$ in the PGF (A.19). Thus, we see that the BDI:I_0 process can be expressed as a sum of two statistically independent processes:

$$I_{BDI:I_0}(t) = I_{BDI:I_0=0}(t) + I_{BD:I_0}(z, t). \tag{57}$$

It can be further shown that the the birth-and-death process can be decomposed into two processes: the **the pure-birth process** and a generalized binomial distributed process, denoted $I_{GB:p(t)}(t)$, where $p(t)$ is given by (87). Thus,

$$I_{BD:I_0}(t) = I_{NB(I_0, \beta(t))}(t) + I_{GB:I_0}(t), \tag{58}$$

and the PGFs of these processes can be expressed as

$$G_{BD:I_0}(z, t) = G_{NB(I_0, \beta(t))}(z, t) \ G_{GB:I_0}(z, t) \tag{59}$$

where

$$G_{NB(I_0, \beta(t))}(z, t) = \left(\frac{1 - b(t)}{1 - b(t)z} \right)^{I_0} \tag{60}$$

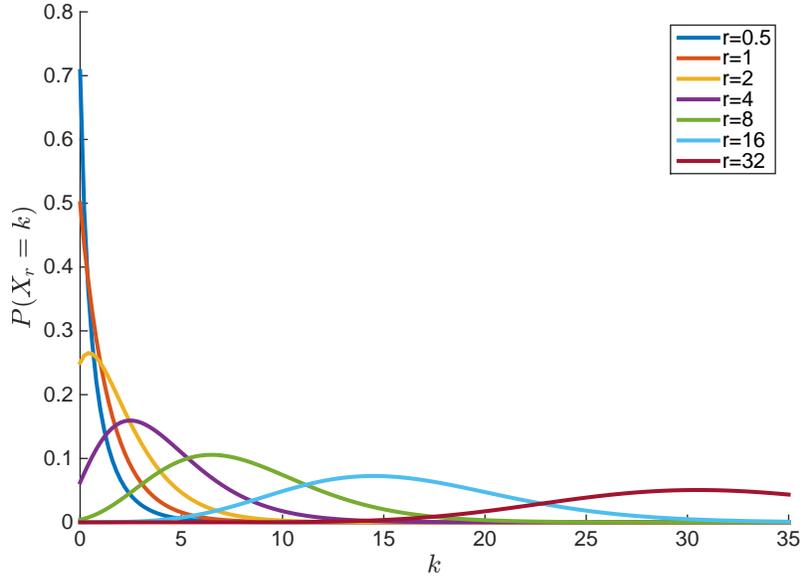$$G_{GB:p(t)}(z, t) = (1 - p(t) + p(z)z)^{I_0} \tag{61}$$

Figure 6: The (shifted) negative binomial distribution (39), i.e., the probability distribution of the number of failures needed to achieve r successes in Bernoulli trials with $q(=1-p)=0.5$ and for $r=0.5, 1, 2, 4, 8, 16, 32$.:

5 Important Properties of the Negative Binomial Distribution (NBD)

5.1 Coefficient-of-variation of an NBD with small r

Let us further examine properties of a random variable X of $NB(r, q)$ defined by its PMF (39) and PGF (41). Its mean and variance are given in (42). The mode of the this distribution is given by

$$\text{mode}_{X_{NB}} = \begin{cases} \lfloor \frac{q(r-1)}{1-q} \rfloor, & \text{if } r > 1 \\ 0, & \text{if } r \leq 1. \end{cases} \tag{62}$$

Recall that the parameter r in our problem is the ratio of ν [person/unit-time] (the rate with which an infected person arrives in the community in question) to λ [person/unit-time/person] (the rate with which an infectious person infects susceptible people. As long as tight security measures are enforced at the boundaries of the community, ν is kept small. Thus, in our problem of practical interest, $r < 1$, or even $r \ll 1$.

Figure 6 shows the (shifted) negative binomial distribution (NBD) for $p = 1 - q = 0.5$ and for various values of $r = 0.5, 1, 2, 4, 8, 16, 32$.

As we can see in Figure 6, the PMF with $r = 0.5$ falls off rapidly by around $n \approx 5$. For $q = 0.5$ and $r = 0.5$, we compute the first several values of P_n:

$$P_0 = 0.7071, \quad P_1 = 0.5303, \quad P_2 = 0.3315, \quad P_3 = 0.1934, \quad P_4 = 0.1088, \quad P_5 = 0.0598, \cdots, P_{10} = 0.00256. \tag{63}$$

In order to see more clearly how the PMF P_n looks for small r, let us rewrite (39) as follows:

$$P_n = k(r) \frac{(n-1+r)}{(n-1)} \frac{(n-2+r)}{(n-2)} \cdots \frac{(2+r)}{2} \frac{(1+r)}{1} \frac{q^n}{n}, \quad \text{where} \quad k(r) = rp^r. \tag{64}$$

The leading term $k(r)$ does not depend on n. So we can write

$$P_n \propto \frac{(n-1+r)}{(n-1)} \frac{(n-2+r)}{(n-2)} \cdots \frac{(2+r)}{2} \frac{(1+r)}{1} \frac{q^n}{n}. \tag{65}$$

In the limit $r \to 0$, the shape of P_n will become

$$P_n \propto \frac{q^n}{n}. \tag{66}$$

Recall the following Maclaurin series expansion

$$-\ln(1-q) = q + \frac{q^2}{2} + \frac{q^3}{3} + \cdots, \tag{67}$$

from which we obtain

$$\sum_{n=1}^{\infty} \frac{-1}{\ln(1-q)} \frac{q^n}{n} = 1. \tag{68}$$

Thus, we find (66) can be written as

$$P_n = \frac{1}{-\ln(1-q)} \frac{q^n}{n}, \quad n = 1, 2, 3, \cdots. \tag{69}$$

This distribution is called the **logarithmic distribution** (a.k.a. **logarithmic series distribution**), which was introduced by R. A. Fisher in 1943 [25]. The PGF of (69) is given, using (67) once again, by

$$G_{\log}(z) = \frac{\ln(1-qz)}{\ln(1-q)}. \tag{70}$$

In our case the probability q is given by $\beta(t)$ of (45):

$$q = \beta(t) = \frac{\lambda(e^{at} - 1)}{\lambda e^{at} - \mu}, \quad \text{where} \quad a = \lambda - \mu. \tag{71}$$

Thus, for sufficiently large t, the parameter q is very close to one, thus the term q^n is a slowly decreasing geometric series. For small r, the coefficient is also slowly decaying. Thus, for small r and q close to 1, the NBD is a very slowly decaying distribution. having a long tail, similar to **Zipf's law** or the **zeta distribution** with small power exponent α (see e.g., [15], pp. 62-64).

Example 2: Consider the same set of model parameters used in Example 1, i.e., $\lambda = 0.3, \mu = 0.1$ and $\nu = 0.2$. Then $r = \frac{\nu}{\lambda} = \frac{2}{3}$, and $a = \lambda - \mu = 0.2$. In Figure 7 through Figure 12, we show the PMF $P_n(t)$ (47) of the infected population $I(t)$ at $t = 1, 5, 10, 20, 30$ and 50. In all these cases, $I_0 = 0$, i.e., there are no infected individuals at $t = 0$.

One useful method to see clearly the tail end of the distribution function is to plot the **log-survival function**, $\log_{10}(1 - F_n)$, where F_n is the cumulative distribution function (CDF), i.e., $F_n = \sum_{i=0}^{n} P_n(t)$, and its complement $1 - F_n$ is called the survival (or survivor) function.[15] In Figure 13 and Figure 14, we show the log-survivor functions of the distribution of $I(t)$ at $t = 10$ and $t = 30$. We confirm our earlier observation that the NBD behaves similar to a geometric distribution, which is a straight line in the log-survivor plot.

---

[15]The term "survival (or survivor) function" is often used in reliability theory. If a continuous variable X represent the "life" of a human, or a product (e.g., electric bulb), with its probability density function (PDF) $f_X(x)$ and cumulative distribution function (CDF) $F_X(x) = \int_0^t f_X(s)\, ds$, the survival function is defined by $1 - F_X(x)$. By plotting $\log(1 - F_X(x)$ vs. x, we find more clearly, how the remaining life will behave than a regular plot of $f_X(x)$ or $F_X(x)$. See e.g., [15] pp. 144-146.
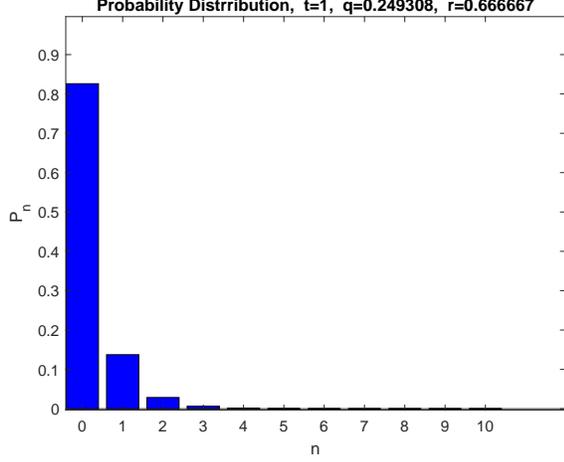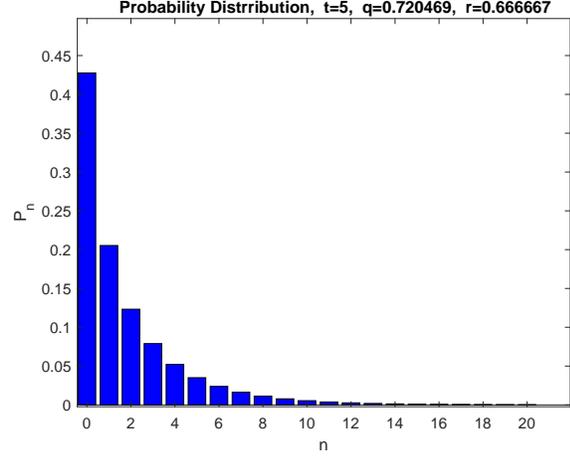
Figure 7: $P_n(t)$ at $t = 1$.
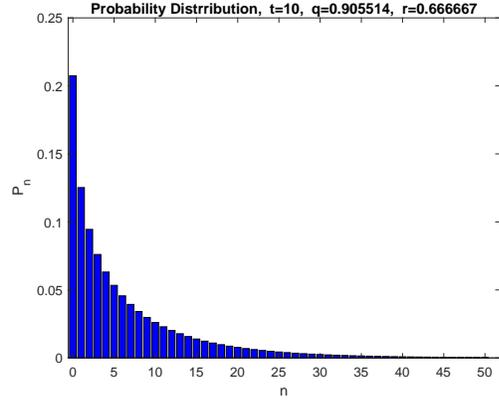

Figure 8: $P_n(t)$ at $t = 5$.
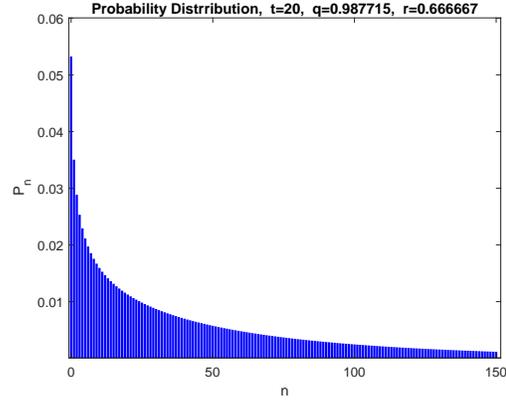

Figure 9: $P_n(t)$ at $t = 10$.

Figure 10: $P_n(t)$ at $t = 20$.

As we observed in Example 5.1, the PMF $P_n(t)$ looks similar to a slowly decaying geometric distribution for modest values of t, with a long tail, which gives a large variance. An often used measure of dispersion is the **coefficient of variation** (**CV**). From (53) we can derive the following expression:

$$\sigma^2_{BDI}(t) = \overline{I}^2(t)\frac{\lambda e^{at} - \mu}{\nu(e^{at} - 1)} = \frac{\overline{I}^2(t)}{r\beta(t)}, \tag{72}$$

from which we find the CV at time t of the BDI process is given by

$$c_{BDI}(t) = \frac{\sigma_{BDI}(t)}{\overline{I}(t)} = \frac{1}{\sqrt{r\beta(t)}}. \tag{73}$$

It is evident that $\beta(t)$ quickly approaches unity from below, as t increases[16]

$$\beta(t) \approx 1, \quad \text{for} \quad t \gg 1, \quad \text{and} \quad \lim_{t\to\infty} \beta(t) = 1. \tag{74}$$

Thus, we have arrive at the following proposition regarding the CV of the BDI process:

---

[16]When t is such $at = 7$, for instance, will make $\beta(t)$ within around 2% off from the unity. In our example of $a = 0.2$[/day], $t = 23$[days] will make $\beta(t)$ practically equal to unity.
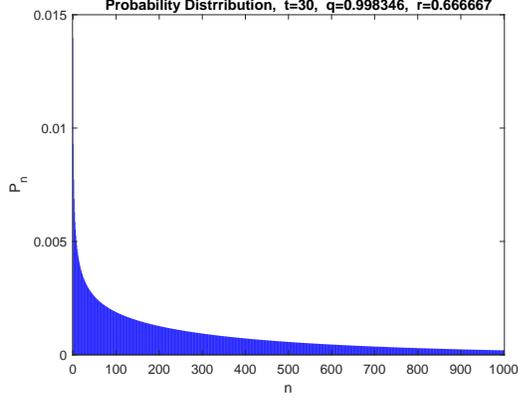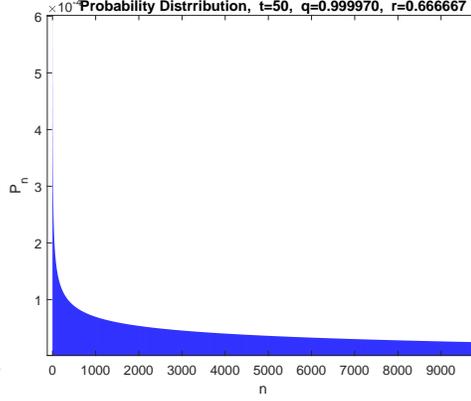
Figure 11: $P_n(t)$ at $t = 30$.
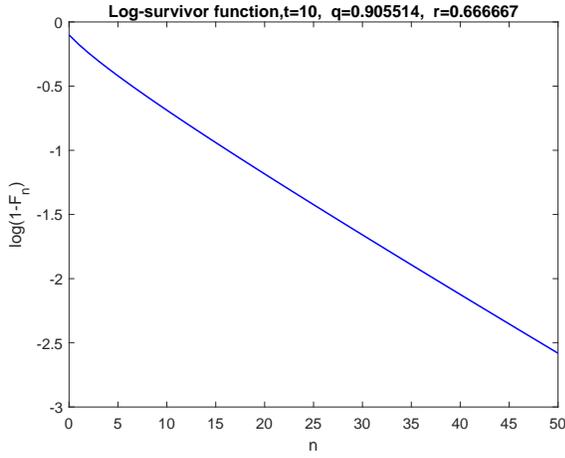

Figure 12: $P_n(t)$ at $t = 50$.

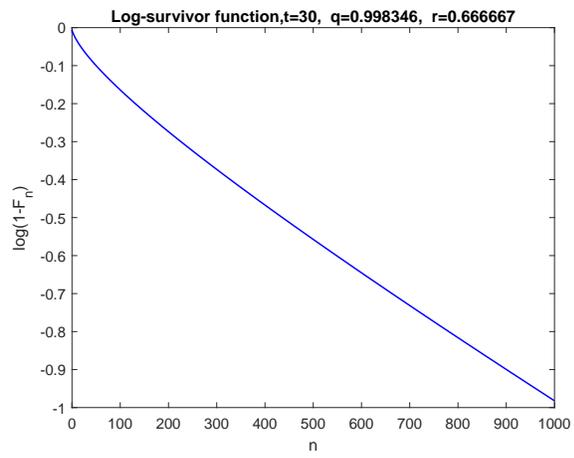Figure 13: The log-survivor function plot of the distribution of $I(t)$ at $t = 10$.

Figure 14: The log-survivor function plot of the distribution of $I(t)$ at $t = 30$.

Proposition 1. The coefficient of variation *of the BDI process rapidly converges to a constant* $\sqrt{r^{-1}}$ *for all t such that* $(\lambda - \mu)t \geq 7$:

$$\lim_{t \to \infty} c_{BDI}(t) = \sqrt{r^{-1}}, \quad where \quad r = \frac{\nu}{\lambda}. \tag{75}$$

Example 3:

Consider the environment considered in Example 3.2, i.e., $\nu = 0.2$/day, $\lambda = 0.3$/day/infectious person, $\mu = 0.1$/day/infectious person. Thus, $R_0 = \lambda/mu = 3$, $a = \lambda - \mu = 0.2$., and $r = \frac{\nu}{\lambda} = 2/3$. Hence, $\sqrt{r^{-1}} = 1.225$.

The mean, variance and the standard deviation of $I(t)$ on the days $t = 30$ and 50 are:

$\overline{I}(0) = 0,$	$\mathrm{Var}[I(0)] = 0,$	$\sigma_I(0) = 0,$	$c_I(0) = \text{undefined},$
$\overline{I}(30) = 402.5,$	$\mathrm{Var}[N(30)] = 243,411.9,$	$\sigma_N(30) = 493.4,$	$c_N(30) = 1.226,$
$\overline{I}(50) = 22,025.5,$	$\mathrm{Var}[N(50)] = 727,706,000.9,$	$\sigma_N(50) = 26,976.0,$	$c_N(50) = 1.225.$

So, we have confirmed that the coefficient of variation is bounded from below by $\sqrt{r^{-1}} = 1.225$. $\square$

19

5.2 Relation between the NBD and "generalized" binomial distributions

In the previous section we gave a formal definition of the *generalized* negative binomial distribution, by allowing the parameter r to be a positive real number. We now want to generalize the binomial distribution $B(N, p)$, by allowing N to be a real number (positive or negative), and p be any real number, positive or negative. Thus, we cannot assign any probabilistic interpretation.

Definition 3. *Consider a generating function of z, which takes the form for real numbers, p and α.*

$$G(z) = (1 - p + p\, z)^\alpha. \tag{76}$$

Then its inverse transform

$$f_n = \binom{\alpha}{n} p^n (1 - p)^{\alpha - n}, \quad n = 0, 1, 2, \ldots \tag{77}$$

*is called a **generalized binomial distribution**, denoted GB(n, p).*

There is no guarantee that each f_n non-negative, let alone $0 \le f_n \le 1$, However, they add up to unity, because $G(1) = 1$. The main reason why we wish to generalize the binomial distribution is for the convenience of computing $P_n(t)$ for the BDI process and other birth-and-death processes (without immigration) discussed.

Proposition 2 (A generalized binomial distribution associated with a (generalized) negative binomial distribution). *For a given (generalized) negative binomial distribution, whose PGF is given by*

$$G_{NB}(z) = \left(\frac{1 - q}{1 - qz} \right)^r, \tag{78}$$

with the PMF

$$P_n^{NB} = \binom{r + n - 1}{n} (1 - q)^r q^n, \quad n = 0, 1, 2, \cdots, \tag{79}$$

its associated generalized binomial distribution is given by

$$f_n = \binom{-r}{n} p^n (1 - p)^{-r - n} \tag{80}$$

where p and q satisfy the relation

$$(1 - p)(1 - q) = 1, \quad or \quad p = -\frac{q}{1 - q}. \tag{81}$$

Proof. First we compare the generalized binomial coefficients

$$\binom{-r}{n} = \frac{(-r)(-r - 1) \cdots (-r - n + 2)(-r - n + 1)}{n!}$$

$$= \frac{(r + n - 1)(r + n - 2) \cdots (r + 1)r(-1)^n}{n!} = (-1)^n \binom{r + n - 1}{n} \tag{82}$$

We also find

$$p^n (1 - p)^{-r - n} = \left(\frac{-q}{1 - q} \right)^n (1 - q)^{r + n} = (-1)^n q^n (1 - q)^r \tag{83}$$

Thus, we have shown the equivalence between (79) and (80).

An alternative way to prove this equivalence is to write the PGF (78) using p. defined by (81):

$$\left(\frac{1-q}{1-qz}\right)^r = (1 - p + p\,z)^{-r}. \tag{84}$$

Recall **Newton's generalized binomial formula**:

$$(1+t)^a = 1 + \binom{a}{1}t + \binom{a}{2}t^2 + \cdots = \sum_{j=0}^{\infty}\binom{a}{j}t^j. \tag{85}$$

Apply this formula to the RHS of (84), by setting $a = -r$ and $t = \frac{pz}{1-p}$. The coefficient of the z^n term is given by f_n of (77) $\square$ $\square$

Now let us revisit the PGF (44), which we can write as

$$G(z,t) = (1 - p(t) + p(t)\,z)^{-r}, \tag{86}$$

where

$$p(t) = -\frac{\lambda e^{at} - \mu}{a}. \tag{87}$$

By referring to Proposition 2 we find that $p(t)$ and $\beta(t)$ are related by

$$p(t) = -\frac{\beta(t)}{1 - \beta(t)}. \tag{88}$$

Thus, by applying Newton's formula, we can obtain the following expression for $P_n(t)$:

$$P_n(t) = \binom{-r}{n}(1 - p(t))^n p(t)^{-r-n}, \quad n = 0, 1, 2, \cdots. \tag{89}$$

Using the binomial coefficient formula and the equivalence (83), i.e.,

$$p^n(1-p)^{-r-n} = (-1)^n \beta^n(1-\beta)^r, \tag{90}$$

we see that $P_n(t)$ given by (89) is equivalent to (47).

5.3   The NBD as a Compound Poisson Distribution

In this section we will investigate how to represent the BDI process $I_{BDI}(t)$ as a compound Poisson process. The negative binomial distribution is **infinitely divisible** (see e.g., Feller [9]), and hence can be represented as a compound Poisson process. First we show this representation by considering a special case of the BDI process where no death occurs.

5.3.1   Birth-and-immigration (BI) process

Let us consider a process which has not been discussed, to the best of our knowledge, in the literature. That is, the pure birth process, accompanied by immigration. Let us call this process a **birth-and-immigration process**, a BI process for short.

The PGF of this process can be readily found by setting $\mu = 0$ in (A.19), and after some manipulation, we obtain

$$G_{BI}(z,t) = \left(\frac{1}{e^{\lambda t} - (e^{\lambda t}-1)z}\right)^r \left(\frac{z}{e^{\lambda t} - (e^{\lambda t}-1)z}\right)^{I_0} = z^{I_0}\left(\frac{1 - \beta_0(t)}{1 - \beta_0(t)z}\right)^{r+I_0}, \tag{91}$$

where

$$\beta_0(t) = 1 - e^{-\lambda t}. \tag{92}$$

Not surprisingly, this $\beta_0(t)$ can be obtained by setting $\mu = 0$ in $\beta(t)$ of (45). Thus, the BI process $I_{BI}(t)$ is distributed according to the negative binomial distribution $\mathrm{NB}(r + I_0, \beta_0(t))$ shifted by I_0. Hence, we find the PMF at time t as given below:

$$P_n^{BI}(t) = e^{-(r+I_0)\lambda t} \binom{n - I_0 + r - 1}{n - I_0} \left(1 - e^{-\lambda t}\right)^{n - I_0}, \quad n = I_0, I_0 + 1, I_0 + 2, \cdots. \tag{93}$$

If, $I_0 = 0$, the PMF reduces to $\mathrm{NB}(r, \beta_0(t))$. By noting $r\lambda = \nu$, we find the first few PMF values as follows:

$$P_0^{BI:0}(t) = e^{-\nu t},$$
$$P_1^{BI:0}(t) = e^{-\nu t} r \left(1 - e^{-\lambda t}\right),$$
$$P_2^{BI:0}(t) = e^{-\nu t} r \frac{(r + 1)}{2!} \left(1 - e^{-\lambda t}\right)^2,$$
$$P_3^{BI:0}(t) = e^{-\nu t} r \frac{(r + 2)(r + 1)}{3!} \left(1 - e^{-\lambda t}\right)^3,$$

where $P_0^{BI:0}(t)$ represents the **idle period** of the system which ends upon the Poisson arrival. This initial idle period is the only time when the system is empty. Since there is no death or departure of any kind, the population monotonically increases as the time elapses.

As we defined in Definition 1 of Section 3.2, $A(t)$ is the cumulative count of infected and infectious persons, arriving from the outside up to time t. Let us count them as $c = 1, 2, \cdots, A(t)$. Consider a "countable attribute" associated with each arrival and denote it N_c. Assume that the N_c's are independent identically distributed (i.i.d.) RVs. Kobayashi and Konheim [26] discuss use of a *compound process* in the context of data communication system, where each arrival of a data packet carries N_c [data units], e.g., [bytes]. Then, the sum

$$S(t) = \sum_{c=1}^{A(t)} N_c \tag{94}$$

should be a quantity of interest in determining, e.g., the buffer size of a statistical multiplexer needed to keep the probability of buffer overflow below some prescribed level.

In our problem at hand, each infected arrival from the outside becomes an "ancestor" who produces many "descendants" over multiple generations, who are all "infected" persons and constitute the members of the "c clan." Such a process is called a **branching process**. In our model, each infected individual, whether an ancestor or a descendant, produces a new descendant at rate λ [person/unit-time].

Let $A_j(t)$ be the number of arrivals in $(0, t]$ and having j *descendants* at time t. The case $j = 0$ does not exist, since there is no death, hence each clan includes at least its ancestor, such that

$$\sum_{j=1}^{\infty} A_j(t) = A(t), \quad t \geq 0. \tag{95}$$

Branson [27], who ascribes the original idea to Karin and McGregor [28], shows that the RVs $A_j(t)$'s are independent Poisson variables with mean $m_j(t)$ where

$$m_j(t) = E[A_j(t)] = \frac{\nu}{\lambda} \frac{\xi^j}{j}, \tag{96}$$

where ξ is a function of λ and t.

The above formula (96) leads to the idea that the probability distribution that each ancestor should produce j descendants (i.e., secondary, tertiary infectees, etc.) whose probability distribution takes the form

$$\Pr[N_c(t) = j] = P_j^{N_c}(t) \propto \frac{\beta_0(t)^j}{j}, \quad j \geq 1. \tag{97}$$

This distribution is the *logarithmic distribution* of (69). The PGF of $N_c(t)$ is therefore given by

$$G_{N_c}(z) = \frac{\ln(1 - \beta_0(t) z)}{\ln(1 - \beta_0(t))} = \frac{\ln(1 - (1 - e^{-\lambda t})z)}{-\lambda t} \tag{98}$$

Note that $N_c(t)$ is the number of the family members at time t of the cth clan, with the ancestor itself included. Since we are assuming no death, the sum of these numbers over all $A(t)$ ancestors should be the total population $I(t)$ at time t:

$$I(t) = \sum_{c=1}^{A(t)} N_c(t). \tag{99}$$

Then, the PGF of $I(t)$ can be expressed as

$$G(z, t) = \mathsf{E}\left[z^{I(t)}\right] = \mathsf{E}\left[z^{\sum_{c=1}^{A(t)} N_c(t)}\right] = \mathsf{E}\left[\left(\mathsf{E}[z^{N_c}]\right)^{A(t)}\right] = \mathsf{E}\left[(G_{N_c}(z, t))^{A(t)}\right]$$

$$= G_A(G_{N_c}(z)) \tag{100}$$

The PGF of the Poisson arrival process of rate ν is given by

$$G_A(z, t) = \exp(\nu t(z - 1)). \tag{101}$$

Thus, we find the PGF of the BI process, a compound Poisson process, is

$$G(z, t) = G_A(G_{N_c}(z)) = \exp\left\{\nu t\left(\frac{\ln(1 - \beta_0(t))}{\ln(1 - \beta_0(t) z)} - 1\right)\right\}$$

$$= \left(\frac{1 - \beta_0(t)}{1 - \beta_0(t) z}\right)^r, \quad \text{where} \quad r = \frac{\nu}{\lambda}, \tag{102}$$

which is the PGF (91), where $I_0 = 0$. Thus, we have shown that the Poisson process with rate ν ends up with the negative binomial distributed process $\mathrm{NB}(r, \beta_0(t))$. It is interesting that the infection process which forms a **branching process** where an infectious person infects a susceptible person at rate λ [persons/time unit] will result in the negative binomial distribution. The *positive feedback loop* of creating its descendants exhibits the negative binomial distributed process.

5.3.2 Birth-and-death with immigration (BDI) process as a compound Poisson process

Now we extend our argument of the previous section to the BDI process with the initial condition $I_0 = 0$, whose PGF is given by (44).

$$G(z, t) = \left(\frac{1 - \beta(t)}{1 - \beta(t)z}\right)^r, \tag{103}$$

where $\beta(t)$ is defined by (45):

$$\beta(t) = \frac{\lambda(e^{at} - 1)}{\lambda e^{at} - \mu}. \tag{104}$$

The first few PMF values can be found from (47) as

$$P_0^{BDI:0}(t) = B(t), \quad \text{where} \quad B(t) = (1 - \beta(t))^r = \left(\frac{a}{\lambda e^{at} - \mu} \right)^r,$$

$$P_1^{BDI:0}(t) = B(t) r \beta(t),$$

$$P_2^{BDI:0}(t) = B(t) r \frac{(r+1)}{2!} \beta(t)^2,$$

$$P_3^{BDI:0}(t) = B(t) r \frac{(r+2)(r+1)}{3!} \beta(t)^3. \tag{105}$$

The function $\beta(t)$ is between 0 and 1, and approaches 1 as $t \to \infty$. If $\nu \to 0$, then $r \to 0$ as well and $B(t) \to 1$. Since $B(t)r$ does not depend on n, we see again the **Fisher series**:

$$\Pr[N_c(t) = j] = P_j^{N_c}(t) \propto \frac{\beta(t)^j}{j}, \quad j \geq 1. \tag{106}$$

Then by replacing $\beta_0(t) = 1 - e^{-\lambda t}$ by $\beta(t) = \frac{\lambda(e^{at}-1)}{\lambda e^{at}-\mu}$ the entire mathematical steps of the BI process case carry over to the case of the BDI process, except for the last step (102), where the parameter r needs to be defined as

$$r = \frac{-\nu t}{\ln \left(\frac{\lambda - \mu}{\lambda e^{at} - \mu} \right)}. \tag{107}$$

If $a = \lambda - \mu > 0$, then for $t \gg 1$, the following approximation holds

$$r \approx \frac{-\nu t}{-at} = \frac{\nu}{\lambda} \frac{\lambda}{\lambda - \mu}. \tag{108}$$

Therefore, it seems clear that we need to to choose a Poisson process other than the one with mean νt. Let $\nu(t)$ be the mean of the Poisson process which is to be compounded with $N_c(t)$ whose PGF is

$$G_{BDI:0}(t) = \left(\frac{1 - \beta(t)}{1 - \beta(t) z} \right)^r. \tag{109}$$

Then it should be clear from the above discussion that $\nu(t)$ must be given by

$$\exp(\nu(t)) = \left(\frac{\lambda e^{at} - \mu}{\lambda - \mu} \right)^r = \exp(art) \left(\frac{\lambda - \mu e^{-at}}{\lambda - \mu} \right)^r$$

$$= \exp(\nu t) \left(\frac{\lambda - \mu e^{-at}}{\lambda - \mu} \right)^r. \tag{110}$$

Thus,

$$\nu(t) = \nu t + r \ln \left(\frac{\lambda - \mu e^{-at}}{\lambda - \mu} \right). \tag{111}$$

Thus, we have shown that the BDI process as a compound Poisson process with rate $\nu(t)$.

6 Concluding Remarks

At the beginning of Section 3 we showed the results of twelve simulation runs of the BDI process to provoke the reader's interest. Hopefully you have found an answer to the first question. How about the second question? If you are an avid reader, you should have figured out the answer by now. In Part II of this report

under preparation [1], we will present more simulation results, which should help the reader find a definitive answer to the second question.

In this report, we focused on the analysis of $I(t)$, the infected population. Although we presented the stochastic means of $B(t)$ (the population of the secondary infections) and $R(t)$ (the recovered/dead population), we defer a full analysis of these processes to Part II.

We will also show, both by analysis and simulation, how the infection process can be controlled by changing the values of λ, μ and ν at certain points in time. Increasing the so-called *social distance* would decrease the value of λ. The value of μ can be increased by improving medical treatments to speed up the recovery process. Another possible option is to reduce the size of the susceptible population significantly by exposing the young and healthy susceptible to the disease so that they become immune to the disease. The quantitative model such as ours should help policy makers and health officials to make judicious choices of these options by assessing the effects and costs of various options available to them. In Part II (or III), we will also discuss how we should estimate the model parameters from empirical data so that the forecast can be made as reliable as possible.

A Derivation of PGF (46)

The partial differential equation (PDE) (15) can be solved for PGF $G(z,t)$ by Lagrange's method with auxiliary differential equations. We write (15) in the following form of a *planar differential equation* (see e.g., [22], [24], [23], [14], pp. 600-603):

$$p\frac{\partial G}{\partial t} + q\frac{\partial G}{\partial z} = r, \tag{A.1}$$

where G denotes $G(z,t)$ defined above, and p, q and r are, in general, functions of t, z, and G.

The solution $G = f(t,z)$ represents a surface in the $t - z - G$ space. The normal vector to the surface at any point (t_0, z_0, G_0) is perpendicular to the line through this point, whose direction numbers are the values of p, q and r evaluated at this point, which we denote by p_0, q_0 and G_0. This line has the equations

$$\frac{t - t_0}{p_0} = \frac{z - z_0}{q_0} = \frac{G - G_0}{r_0}. \tag{A.2}$$

Hence, at each point (p, q, G) of the surface, there is a normal vector whose direction numbers, dt, dz and dG, satisfy the following equation:

$$\frac{dt}{p} = \frac{dz}{q} = \frac{dG}{r}. \tag{A.3}$$

In our problem at hand, we set

$$p = 1, \tag{A.4}$$
$$q = -(\lambda z - \mu)(z - 1), \tag{A.5}$$
$$r = \nu(z - 1)G, \tag{A.6}$$

Then (A.3) becomes

$$\frac{dt}{1} = -\frac{dz}{(\lambda z - \mu)(z - 1)} = \frac{dG(z,t)}{\nu(z - 1)G(z,t)}. \tag{A.7}$$

which leads to the following two independent solutions:

$$\left(\frac{z - 1}{\mu - \lambda z}\right) e^{at} = C_1, \text{ where } a = \lambda - \mu, \tag{A.8}$$

and

$$G(z,t)(\mu - \lambda z)^r = C_2, \quad \text{where} \quad r = \frac{\nu}{\lambda} \tag{A.9}$$

where C_1 and C_2 are integration constants. We write the functional relations between C_1 and C_2 as

$$C_2 = f(C_1),, \tag{A.10}$$

which, together with (A.7), implies the following general solution:

$$G(z,t) = (\mu - \lambda z)^{-r} f\left(\left(\frac{z-1}{\mu - \lambda z}\right) e^{\lambda t}\right). \tag{A.11}$$

From the initial condition (17), we find

$$G(z,0) = (\mu - \lambda z)^{-r} f\left(\frac{z-1}{\mu - \lambda z}\right) = z^{I_0}. \tag{A.12}$$

from which we find

$$f\left(\frac{z-1}{\mu - \lambda z}\right) = z^{I_0}(\mu - \lambda z)^r. \tag{A.13}$$

We introduce a new variable y by

$$\frac{z-1}{\mu - \lambda z} = y, \quad \text{i.e.,} \quad z = \frac{1+\mu y}{1+\lambda y}. \tag{A.14}$$

From the last two equations, we can determine the functional form of f:

$$f(y) = \left(\frac{1+\mu y}{1+\lambda y}\right)^{I_0} \left(\frac{\mu - \lambda}{1+\lambda y}\right)^r. \tag{A.15}$$

By substituting the last equation into (A.11), we obtain

$$G(z,t) = (\mu - \lambda z)^{-r} f\left(\left(\frac{z-1}{\mu - \lambda z}\right) e^{at}\right) \tag{A.16}$$

By noting

$$\frac{1+\mu y}{1+\lambda} = \frac{\mu - \lambda z + \mu(z-1)e^{at}}{\mu - \lambda z + \lambda(z-1)e^{at}} \tag{A.17}$$

and

$$\frac{\mu - \lambda}{1+\lambda y} = \frac{(\mu - \lambda)(\mu - \lambda z)}{\mu - \lambda z + \lambda(z-1)e^{at}} \tag{A.18}$$

we finally obtain the PGF of $I(t)$

$$\begin{aligned}
G(z,t) &= (\mu - \lambda z)^{-r} \left(\frac{(\mu - \lambda)(\mu - \lambda z)}{\mu - \lambda z + \lambda(z-1)e^{at}}\right)^r \left(\frac{\mu - \lambda z + \mu(z-1)e^{at}}{\mu - \lambda z + \lambda(z-1)e^{at}}\right)^{I_0} \\
&= \left(\frac{\mu - \lambda}{\mu - \lambda z + \lambda(z-1)e^{at}}\right)^r \left(\frac{\mu - \lambda z + \mu(z-1)e^{at}}{\mu - \lambda z + \lambda(z-1)e^{at}}\right)^{I_0}
\end{aligned} \tag{A.19}$$

Note that the PGF is given as a product of two PGFs, the second being the PGF of the birth-and-death process without immigration.

If $I_0 = 0$, i.e., if the system is initially empty, the above PGF reduces to

$$G(z,t) = \left(\frac{a}{\lambda e^{at} - \mu - \lambda(e^{at} - 1) z}\right)^r, \quad \text{when} \quad I(0) = I_0 = 0. \tag{A.20}$$

Acknowledgments

References

[1] H. Kobayashi, "Stochastic Modeling of an Infectious Disease: Part II: Validation by Simulation Experiments (under preparation)." `http://hp.hisashikobayashi.com`, 2020.

[2] W. O. Kermack and A. G. McKendrick, "A Contribution to the Mathematical Theory of Epidemics," *Proc. Royal Soc. London*, vol. A 115, pp. 700–721, 1927.

[3] N. Bacaër, "The model of Kermack and McKendrick for the plague epidemic in Bombay and the type reproduction number with seasonality," *Journals of Mathematical Biology*, vol. 64, pp. 403–422, March 2012.

[4] R. M. Anderson and R. M. May, "Population Biology of Infectious Diseases: Part I," *Nature*, vol. 280, pp. 361–367, August 1979.

[5] R. M. Anderson and R. M. May, *Infectious Diseases of Humans: Dynamics and Control*. Oxford University Press, 1991.

[6] M. Martcheva, *An Introduction to Mathematical Epidemiology*. Springer, 2010.

[7] W. Feller, "Die Grundlagen der Volterraschen Theorie des Kampfes ums Dasein in wahrscheinlichkeitstheoritischer Behandlung," *Acta Biotheoretica*, vol. 5, pp. 11–40, 1939.

[8] D. G. Kendall, "The generalized "birth-and-death" process," *Ann. Math. Statist.*, vol. 19, pp. 1–15, 1948.

[9] W. Feller, *Introduction to Probability and Its Applications: Vol. I*. John Wiley & Sons, 1968.

[10] R. Syski, *Introduction to Congestion Theory in Telephone Systems*. North Holland, 1986.

[11] L. Kleinrock, *Queueing Systems, Vol. I: Theory*. John Wiley & Sons, Inc., 1975.

[12] F. P. Kelly, *Reversibility and Stochastic Networks*. John Wiley & Sons, Inc., 1979.

[13] H. Kobayashi, *Modeling and Analysis: An Introduction to System Performance Evaluation Methodology*. Addison Wesley, 1978.

[14] H. Kobayashi and B. L. Mark, *System Modeling and Analysis: Foundations for System Performance Evaluation*. Prentice Hall, 2008.

[15] H. Kobayashi, B. L. Mark, and W. Turin, *Probability, Random Processes, and Statistical Analysis*. Cambridge University Press, 2012.

[16] H. Takagi, "Lecture Note: Bith-and-Death Process and Its Application (in Japanese)," March 2007.

[17] M. Bartlett, *An Introduction to Stohasctic Processes with Special Reference to Methods and Applications*. Cambridge University Press, 3 ed., 1978.

[18] D. G. Kendall, "On some modes of population growth leading to R. A. Feller's logarithmic series distribution," *Biometrika*, vol. 19, pp. 6–15, 1948.

[19] M. Bartlett, "Some Evolutionary Stochstic Processes," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 11, no. 2, pp. 211–229, 1949.

[20] N. T. Bailey, *The Elements of Stochastic Processes With Applications to the Natural Sciences.* John Wiley & Sons, Inc., 1964.

[21] L. J. D. Allen, *An Introduction to Stochastic Processes with Application to Biology.* Chapman & Hall/CRC, 2 ed., 2011.

[22] D. Gross and C. Harris, *Fundamentals of Queueing Theory.* John Wiley & Sons, 2 ed., 1985.

[23] Q. Ren and H. Kobayashi, "Transient solutions for the buffer behavior in statistical multiplexing," *Performance Evaluation*, vol. 23, pp. 65–87, 1995.

[24] H. Kobayashi and Q. Ren, "A mathematical theory for transient analysis of computer communication networks (invited paper)," *IEICE Trans. Commun.*, vol. E-75B, pp. 1266–1276, 1992.

[25] R. A. Fisher, A. S. Corbet and C. B. Williams, "The relation between the number of species and the number of individuals in a random sample of an animal population," *Journal of Animal Ecology*, vol. 12, no. 1, pp. 42–58, 1943.

[26] H. Kobayashi and A. G. Konheim, "Queuing Models for Computer Communications Systems Analysis (Invited Paper)," *IEEE Trans. on Communications*, vol. 25, pp. 148–157, January 1977.

[27] D. Branson, "In homogeneous birth-death and birth-death-immigration processes and the logarithmic series distribution," *Stochastic Processes and their Applications*, vol. 39, pp. 131–137, 1991.

[28] S. Karlin and J. McGregor, "The number of mutant forms maintained in a population," in *Proc. 5th Berkeley Symp. Math Statist. Probab.* (L.M. Le Cam and J. Neyman, ed.), vol. 4, (Berkeley, CA), pp. 415–438, Univ. California Presss, 1967.