Stochastic Modeling of an Infectious Disease Part III-D: Further Analysis of the Time-Nonhomogeneous BDI Process Model

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

July 16, 2021

Abstract

We present further results concerning the time-varying BDI (birth-death-with-immigration) process that we have been investigating as a probabilistic characterization of an infectious disease such as COVID-19.

In Part III-B [1] we showed that if the ratio $\nu(t)/\lambda(t)$ (where $\nu(t)$ is the arrival rate of infectious individuals from outside and $\lambda(t)$ is the secondary infection rate) is a constant for all time t, the probability distribution of the infected of all the secondary infections is a time varying negative binomial distribution (NBD) at any t. In this article, we analyze how the failure of the constant ratio assumption will affect the distribution by introducing a multiplicative factor $G_C(z,t)$ as a correcting term in the PGF, which shows how the distribution deviates from the the NBD.

In Section 2, we present a closed-form expression for the conditional PGF of the probability that a time-varying BDI process $I(t)$ takes a value i at t, given it has taken value j at u, where $0 \leq u \leq t$. This is a generalization of the new result we presented as Proposition 1 [1].

In forthcoming reports, the conditional mean and variance that can be found from the above PGF will be used to evaluate, by making use of the *saddle-point integration*, an approximate probability density function (PDF) of a continuous-valued process $X(t)$ which will approximate $I(t)$.

Keywords: Time-varying BDI (birth and death with immigration) process; Correcting factor, Conditional PGF, Negative binomial distribution (NBD),

Contents

*The Sherman Fairchild University Professor of Electrical Engineering and Computer Science, Emeritus. Email: Hisashi@Princeton.EDU, Website: `https://hp.hisashikobayashi.com/`, Wikipedia: `https://en.wikipedia.org/wiki/Hisashi_Kobayashi`

1 Effect of $r(t) = \frac{\nu(t)}{\lambda(t)}$ Not Being a Constant

In our previous paper [1] we have obtained a closed form expression for the probability generating function (PGF) of a general time-nonhomogeneous BDI model. The PGF takes the form

$$G_{BDI:I_0}(z,t) = G_{BD:I_0}(z,t)G_{BDI:0}(z,t), \tag{1}$$

where $G_{BD:I_0}(z,t)$ is the PGF of the BD (birth-death) process with the initial population I_0, originally derived by Kendall [2], shown in (70) of [3]

$$G_{BD:I_0}(z,t) = \left(1 + \frac{e^{s(t)}(z-1)}{1 - e^{s(t)}L(t)(z-1)}\right)^{I_0}. \tag{2}$$

where $s(t)$ is defined as (cf, [3] (10))

$$s(t) = \int_0^t (\lambda(u) - \mu(u))\,du, \tag{3}$$

and $L(t)$ is defined (cf. ibid., (77)) as

$$L(t) = \int_0^t \lambda(u)e^{s(u)}\,du. \tag{4}$$

The second component of (1) $G_{BDI:0}(z,t)$ represents the contribution by immigrants and their descendants, which the present author has obtained (see Proposition 1 [1])[1]:

$$\boxed{G_{BDI:0}(z,t) = \exp\left(\int_0^t \frac{\nu(u)e^{s(t)-s(u)}(z-1)}{1 - e^{s(t)}(L(t) - L(u))(z-1)}\,du\right).} \tag{5}$$

We have also shown that this second component can be further decomposed to

$$\boxed{G_{BDI:0}(z,t) = G_{NB}(z,t)G_C(z,t),} \tag{6}$$

where the first component is the PGF of a negative binomial distributed (NBD) process $\mathrm{NB}(r(0), \beta(t))$, where $r(0) = \frac{\nu(0)}{\lambda(0)}$.

$$\boxed{G_{NB}(z,t) \triangleq \frac{1}{\left(1 - e^{s(t)}L(t)(z-1)\right)^{r(0)}} = \left(\frac{1-\beta(t)}{1-\beta(t)z}\right)^{r(0)},} \tag{7}$$

[1] Note that we use the subscript $BDI:0$ instead of $ID:0$ used in [1].

The function $\beta(t)$ is defined in ibid. (77) as

$$\beta(t) \triangleq \frac{L(t)}{1+M(t)}, \tag{8}$$

where $M(t)$ is defined, similar to $L(t)$ of (4), as (see ibid. (60))

$$M(t) = \int_0^t \mu(u) e^{-s(u)}\, du. \tag{9}$$

The second term of (6), $G_C(z,t)$, is defined by (22) of [1]:

$$\boxed{G_C(z,t) \triangleq \exp\left(- \int_0^t r'(u) \log\left(1 - e^{s(t)}(L(t) - L(u))(z-1)\right)\, du \right),} \tag{10}$$

which expresses the correcting effect on $G_{BDI:0}(z,t)$, when $r'(u) \neq 0$. Otherwise, $G_C(z,t) = 1$ for all z and t.

Now we note the following properties of the PGF $G_C(z,t)$:

1. Let us define

$$A(z,t,u) \triangleq 1 - e^{s(t)}(L(t) - L(u))(z-1), \tag{11}$$

which leads to

$$G_C(z,t) \triangleq \exp\left\{ - \int_0^t r'(u) \log A(z,t,u)\, du \right\} \tag{12}$$

we can readily see that at $z = 1$

$$G_C(1,t) = 1, \quad \text{for all} \quad t, \tag{13}$$

which implies that $G(z,t)$ qualifies to be a PGF, if its coefficients of z in a polynomial expansion are all non-negative.

If $r'(u) = 0$, then, as noted earlier, we have

$$G_C(z,t) = 1 \quad \text{for all} \quad z \quad \text{and} \quad t. \tag{14}$$

2. Consider a random process that has this $G_C(z,t)$ as its PGF, and we denote this process by $I_C(t)$:

$$G_C(z,t) = \mathsf{E}[z^{I_C(t)}]. \tag{15}$$

Then, the expected value of such process can be found from the PGF (12) as

$$\mathsf{E}[I_C(t)] = e^{s(t)} \int_0^t r'(u)(L(t) - L(u))\, du. \tag{16}$$

Similarly, we can find the variance as

$$\sigma^2_{I_C}(t) = \mathsf{E}[I_C(t)^2] - \mathsf{E}[I_C(t)]^2 = e^{2s(t)} \int_0^t r'(u)(L(t) - L(u))^2 \, du. \tag{17}$$

The RHSs of both (16) and (17) can be simplified by using integration in part:

$$\begin{aligned}\int_0^t r'(u)(L(t) - L(u))\, du &= [r(u)(L(t) - L(u)]_{u=0}^t + \int_0^t r(u)L'(u)\, du \\ &= \int_0^t \nu(u)e^{-s(u)}\, du - r(0)L(t) = N(t) - r(0)L(t).\end{aligned} \tag{18}$$

$$\begin{aligned}\int_0^t r'(u)(L(t) - L(u))^2\, du &= \left[r(u)(L(t) - L(u))^2\right]_{u=0}^t + 2\int_0^t r(u)L'(u)(L(t) - L(u))e^{-s(u)}\, du \\ &= 2L(t)N(t) - 2\int_0^t \nu(u)L(u)e^{-s(u)}\, du - r(0)L(t)^2\end{aligned} \tag{19}$$

Thus, the c.v. (coefficient of variation) of the correction process $I_C(t)$ is

$$\begin{aligned}c_{I_C}(t) &= \frac{\sqrt{\int_0^t r'(u)(L(t) - L(u))^2\, du}}{\int_0^t r'(u)(L(t) - L(u))\, du} \\ &= \frac{\sqrt{2L(t)N(t) - 2\int_0^t \nu(u)L(u)e^{-s(u)}\, du - r(0)L(t)^2}}{N(t) - r(0)L(t)}\end{aligned} \tag{20}$$

2 State Transition Probabilities and the Probability Generating Function

Up to now, Part I through Part III-B, we have been concerned with the probability distributions and their PGF, when the infection process $I(t)$ begins at time $t = 0$ with a given initial value I_0. But in many situations, we may be interested in the future behavior of the process $I(t)$, given its value at some moment u, where $0 \leq u \leq t$. In order to answer such a question, we need to obtain state transition probabilities of the Markov process $I(t)$ from the current value $I(u) = j$ to its value at arbitrary t, $I(t) = k$ for arbitrary u and any state pairs (j, k) which are both in $\mathcal{Z}_+$, where

$$\mathcal{Z}_+ \triangleq \{0, 1, 2, 3, \cdots\}. \tag{21}$$

Let us define

$$P_{jk}(u, t) \triangleq \mathsf{P}[I(t) = k | I(u) = j], \quad j, k \in \mathcal{Z}_+, \quad 0 \leq u \leq t < \infty \tag{22}$$

and we want to find closed form expressions for this set of probabilities.

In principle, we can derive an answer from the closed form solutions that we obtained in [1], which were found for any initial condition $I_0 \in \mathcal{Z}_+$, but such a translation or interpretation is not so obvious in the time-nonhomogeneous situation and could be prone to mistakes. So we go back to the original PDE and derive the solution, and see how our previous result (i.e., for $u = 0$ and $j = I_0$) will come out of this general case.

Consider an infinitesimal interval $(t - dt, t]$. Then we need to consider the following events that might occur for $I(t)$ to take on state k at t:

i The system was in state $k-1$ and an immigrant arrives (i.e., “an infected person enters from outside” in our context) in the interval $(t-dt, t]$. Such situation happens with probability $P_{j,k-1}(t-dt)\nu(t)dt$.[2]

ii The system was in state $k-1$ at time $t-dt$, and one birth takes place in the interval $(t-dt, t]$. Since each of the $k-1$ individuals found in the system will give a birth (i.e., “infect a susceptible individual” in our context) with rate $\lambda(t)$, this situation occurs with probability $P_{j,k-1}(t-dt)(k-1)\lambda(t)dt$.

iii A death occurs (i.e., “an infected person, recovers, gets removed or dies, hence stop being infectious” in our context), moving the system from state $k+1$ to state k at time t. This situation occurs with probability $P_{j,k+1}(t-dt)(k+1)\mu(t)dt$.

iv The system was in state k at $t-dt$ and no events occur in the interval $(t-dt, dt]$. This has probability $P_{j,k}(t-dt)\left[1-(\nu(t)+(k-1)\lambda(t)+(k+1)\mu(t))\,dt\right]$.

v Multiple events occur with probability of order $o(dt^2)$.

Noting that these events are mutually exclusive, we have the following equation:

$$\begin{aligned} P_{j,k}(u,t) &= P_{j,k-1}(u,t-dt)\left[\nu(t)+(k-1)\lambda(t)\right]dt + P_{j,k+1}(u,t-dt)(k+1)\mu(t)dt \\ &\quad + P_{j,k}(u,t-dt)\left[1-(\nu(t)+k\lambda(t)+k\mu(t))\,dt\right] + o(dt^2), \quad u \le t \end{aligned} \tag{23}$$

from which we obtain the following differential-difference equation, a.k.a. *Kolmogorov's forward equation*:

$$\frac{dP_{j,0}(u,t)}{dt} = \mu(t)P_{j,1}(u,t) - \nu(t)P_{j,0}(u,t), \quad j \ge 0, \quad u \le t \tag{24}$$

$$\begin{aligned} \frac{dP_{j,k}(u,t)}{dt} &= \left[\nu(t)+(k-1)\lambda(t)\right]P_{j,k-1}(u,t) + (k+1)\mu(t)P_{j,k+1}(u,t) \\ &\quad - \left[\nu(t)+k\lambda(t)+k\mu(t)\right]P_{j,k}(u,t), \quad k \ge 1, \quad j \ge 0, \quad u \le t \end{aligned} \tag{25}$$

Then by defining the following PGF of the transition probability

$$G_j(z,u,t) \triangleq \mathsf{E}[z^{I(t)}|I(u)=j] = \sum_{k=0}^{\infty} P_{j,k}(u,t)z^k, \quad u \le t \tag{26}$$

we can transform the countably infinitely many differential equations (24), (25) into one partial differential equation (PDE) for a given $j \in \mathcal{Z}_+$:

$$\frac{\partial G_j(z,u,t)}{\partial t} = (z-1)\left[\lambda(t)z-\mu(t))\frac{\partial G_j(z,u,t)}{\partial z} = \nu(t)G_j(z,u,t)\right], \tag{27}$$

with the boundary condition

$$G_j(z,u,u) = z^j. \tag{28}$$

[2]You may want to write $\nu(t-dt)$ instead, but this does not make any difference in the limit $dt \to 0$, assuming the function $\nu(t)$ is continuous at every t.

Then from the corresponding auxiliary differential equations

$$dt = \frac{dz}{(\lambda(t)z - \mu(t))(z-1)} = \frac{dG_j(z,u,t)}{\nu(t)(z-1)G_j(z,u,t)}, \tag{29}$$

we obtain

$$\frac{dz}{dt} = (z-1)(\mu(t) - \lambda(t)z), \quad u \le t. \tag{30}$$

By changing the variable z to x

$$x = \frac{1}{z-1}, \quad \text{i.e.,} \quad z = 1 + \frac{1}{x}, \tag{31}$$

we obtain

$$\frac{dx}{dt} = a(t)x + \lambda(t), \quad \text{where } \ a(t) = \lambda(t) - \mu(t), \quad u \le t \tag{32}$$

Define a function $s(\tau, t)$ by

$$s(\tau, t) \triangleq \int_\tau^t a(v)\,dv = \int_\tau^t (\lambda(v) - \mu(v))\,dv = s(t) - s(u), \tag{33}$$

where $s(t)$ is defined in (3). We multiply (32) by $e^{-s(u,t)}$, obtaining

$$\frac{d(xe^{-s(u,t)})}{dt} = \lambda(t)e^{-s(u,t)}. \tag{34}$$

By integrating the above w.r.t. t from u to t, we obtain the first solution

$$\frac{e^{-s(u,t)}}{z-1} - L(u,t) = C_1, \quad u \le t, \tag{35}$$

where

$$L(u,t) \triangleq \int_u^t \lambda(\tau)e^{-s(\tau)}\,d\tau = L(t) - L(u). \tag{36}$$

The second solution can be obtained from the first and last terms of the auxiliary equations

$$\frac{dG_j(z,u,t)}{\nu(t)G_j(z,u,t)(z-1)} = dt, \quad u \le t \tag{37}$$

which leads, by following the steps (50)-(52) in [4] (or (3)-(6) in [1]), to

$$C_2 = G_j(z,u,t)\exp\left(-\int_u^t \frac{\nu(\tau)e^{-s(u,\tau)}}{C_1 + L(u,\tau)}\,d\tau\right) \tag{38}$$

On writing the functional relation between the integration constants C_1 and C_2 as

$$C_2 = f(C_1) \tag{39}$$

and setting $t = u$ in (35) and (38), and using the boundary condition (28), we find

$$f(y) = \left(1 + \frac{1}{y}\right)^j. \tag{40}$$

Then from (35), (38), (39) and (40), we obtain, after some algebraic manipulations,

$$G_j(z,u,t) = \exp\left(\int_u^t \frac{\nu(\tau)e^{s(u,t)-s(u,\tau)}(z-1)}{1-e^{s(u,t)}(L(u,t)-L(u,\tau))(z-1)}\,d\tau\right)\cdot\left(1+\frac{e^{s(u,t)}(z-1)}{1-e^{s(u,t)}L(u,t)(z-1)}\right)^j \tag{41}$$

The first term of the above product form corresponds to the PGF contributed by immigrants who arrive after time u and their descendants living at time t, whereas the second PGF component is due to the BD process which has started with the population size j at time u. If we set $u = 0$ and $j = I_0$, these PGFs reduce to (57)-(58) of [4] (or (11)-(12) of [1].)

Thus, we can write the above for any $t \geq u \geq 0$ and for any $j = I(u) \geq 0$

$$G_j(z,u,t) = G_{BDI:I(u)=0}(z,t)\cdot G_{BD:I(u)=j}(z,t) \tag{42}$$

where

$$G_{BDI:I(u)=0}(z,u,t) \triangleq \exp\left(\int_u^t \frac{\nu(\tau)e^{s(u,t)-s(u,\tau)}(z-1)}{1-e^{s(u,t)}(L(u,t)-L(u,\tau))(z-1)}\,d\tau\right) \tag{43}$$

and

$$G_{BD:I(u)=j}(z,u,t) \triangleq \left(1+\frac{e^{s(u,t)}(z-1)}{1-e^{s(u,t)}L(u,t)(z-1)}\right)^j, \tag{44}$$

which is a generalization of (57)-(58) of [4].

By defining

$$\alpha(u,t) \triangleq \frac{M(u,t)}{1+M(u,t)}, \tag{45}$$

$$\beta(u,t) \triangleq \frac{L(u,t)}{1+M(u,t)}, \tag{46}$$

where $M(u,t)$ is defined, similar to $L(u,t)$ of (36), by

$$M(u,t) \triangleq \int_u^t \lambda(\tau)e^{-s(\tau)}\,d\tau = M(t) - M(u). \tag{47}$$

Then we can write

$$G_{BD:I(u)=j}(z,u,t) = \left(\frac{\alpha(u,t)+(1-\alpha(u,t)-\beta(u,t))z}{1-\beta(u,t)z}\right)^j, \quad u \leq t. \tag{48}$$

Similarly, we can write

$$G_{BDI:I(u)=0}(z,u,t) = \frac{G_C(z,u,t)}{\left(1-e^{s(u,t)}L(u,t)(z-1)\right)^{r(u)}}$$
$$= G_C(z,u,t)\left(\frac{1-\beta(u,t)}{1-\beta(u,t)z}\right)^{r(u)}, \quad u \le t, \tag{49}$$

where

$$r(u) = \frac{\nu(u)}{\lambda(u)}, \tag{50}$$

and

$$G_C(z,u,t) \triangleq \exp\left(-\int_u^t r'(\tau)\log\left\{1-e^{s(u,t)}(L(u,t)-L(u,\tau))(z-1)\right\}d\tau\right), \quad u \le t. \tag{51}$$

3 Discussion and Future Plans

1. Converting the PGF obtained above to the conditional PMFs $P_{jk}(u,t)$ requires the same computational procedure as we discussed in the previous report: see [4] (21) (or [3], (76), (79)). When $j \gg 1$, which corresponds to I_0 in [1, 3, 4], the computations to find exact PMFs $P_n(t)$ will be prohibitive.

   But the conditional expectation and variance can be obtained from the above PGF, by following the same steps discussed in [4], (39), (60) and Appendix A. For large $j = I(u)$ and/or sufficiently large t, we know that due to the *central limit theorem* this conditional PMFs should converge in distribution to the normal distribution with the above conditional mean and variance.

2. We are currently pursuing to obtain an approximate probability density function (PDF) $f_X(x)$ of a continuous random variable (RV) X, which approximates the PMF of the integer-valued RV $I(t)$, by using the *saddle-point integration* method, and will report it shortly [5, 6]. We will then have a better understanding how large the initial value j and/or $t-u$ must be in order to approximate the distribution by a normal distribution.

References

[1] H. Kobayashi, "Stochastic Modeling of an Infectious Disease: Part III-B: Analysis of the Time-Nonhomogenous BDI Process, and Simulation Experiments of both BD and BDI Process." `https://arxiv.org/abs/2104.00529` and `http://hp.hisashikobayashi.com`, March 30 2021.

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

[3] H. Kobayashi, "Stochastic Modeling of an Infectious Disease: Part III-A: Analysis of Time-Nonhomogeneous Models." `https://arxiv.org/abs/2101.09109`, and `http://hp.hisashikobayashi.com`, January 19 2021.

[4] H. Kobayashi, "Stochastic Modeling of an Infectious Disease: Part III-C: Recapitulation of Parts I, II, III-A and III-B." http://hp.hisashikobayashi.com, July 14 2021.

[5] H. Kobayashi, "Stochastic Modeling of an Infectious Disease: Part VI-A: Complex Analysis Primer." http://hp.hisashikobayashi.com, (in preparation) 2021.

[6] H. Kobayashi, "Stochastic Modeling of an Infectious Disease: Part VI-B: Approximate Analysis based on Saddle-Point Integration." http://hp.hisashikobayashi.com, (in preparation) 2021.