Numerical solution of a Gamma - integral equation using a higher order composite Newton-Cotes formulas

Aubain H Nzokem
Department of Mathematics & Statistics, York University, Toronto, Canada
E-mail: hilaire77@gmail.com

Abstract. The paper aims at solving a complex equation with Gamma - integral. The solution is the infected size \( p \) at equilibrium. The approaches are both numerical and analytical methods. As a numerical method, the higher-order composite Newton-Cotes formula is developed and implemented. The results show that the infected size \( p \) increases along with the shape parameter \( k \). But the increase has two phases: an increasing rate phase and a decreasing rate phase; both phases can be explained by the instantaneous death rate characteristics of the Gamma distribution hazard function. As an analytical method, the Extreme Value Theory consolidates the numerical solutions of the infected size \( p \) when \( k \geq 1 \) and provides a solution limit \( p = 1 - \frac{1}{2} R \) as \( k \) goes to \( +\infty \).

Keywords: Extreme value theory, Gamma distribution, Newton-Cotes formula.

1. Introduction
In the paper, we consider the equation (1), where the unknown \( p \) is the infected size.

\[
\int_{0}^{\infty} \left( \frac{(-z + k)\Gamma(k, z) + z^ke^{-z}}{k\Gamma(k)} \right)^{pM} dz - \frac{k + 1}{2Rp(1-p)M} = 0 \tag{1}
\]

The equation (1) describes the proportion \( p \) of the population infected at the equilibrium in a SIS model. The Susceptible-Infected-Susceptible (SIS) model is defined within a population of constant size \( M \). The disease spreads within the population is modeled by a Poisson process with a rate \( \lambda I = \beta p(1-p)M \), where \( \beta \) is the transmission rate. The analysis is focused on the disease outbreak, where the reproduction number \( R \) is greater than one. For more details on SIS model and equation (1), see [1, 2].

The life span of each individual in the population is modeled by a Gamma distribution with shape \( k \) and scale \( \lambda \) parameters. In the SIS model, the lifetime of each individual is modeled by the exponential distribution, which is nicely tractable. For related work, using the exponential distribution in SIS model, see [3, 4, 5, 6, 7, 8, 9, 10].

The paper uses a broader class of distribution, Gamma distribution with shape \( k \) and scale \( \lambda \) parameters, to model the life span of each individual.

The Gamma distribution provides [11, 12] a wide range of instantaneous death rates within the population as shown its hazard function in Figure 1. When \( 0 < k < 1 \), the instantaneous...
death rate is very high at birth and decreases with the age over the lifetime. whereas for \( k > 1 \), the instantaneous death rate starts at zero at birth and increases with age as shown Figure 1. For \( k=1 \), we have a classical case of the exponential distribution, as shown in Figure 1, the instantaneous death rate is constant \( (\lambda) \) during the lifetime of each individual. In order to solve equation (1), both numerical and analytical methods will be explored. As numerical method, the composite Newton-Cotes quadrature formulas will be implemented in order to provide an accurate estimation of the infected size \( (p) \). As analytical method, the Extreme Value Theory will be used to determine analytical solution.

![Figure 1. Hazard function \( h_X(x, \lambda, k) \) of the Gamma distribution](image)

The paper is structured as follows; in the next section, we will develop the higher-order composite Newton-Cotes quadrature formula and implemented to compute infected size \( (p) \); and the last section will focus on the analytical solution and analyzing the asymptotic infected size \( (p) \).

2. Composite Newton-Cotes Quadrature Formulas

The Newton-Cotes rules value the integrand \( f \) at equally spaced points \( x_i \) over the interval \([a, b]\); where \( x_i = a + i\frac{b-a}{n} = a + ih \) with \( h = \frac{b-a}{n} \); \( n = Qn_0 \) and \( x_{Qp+Q} = x_{Q(p+1)} \) where \( Q \) is the number of \( h \) within the subinterval \([x_{Qp}, x_{Qp+Q}]\) of interval \([a, b]\).

2.1. Composite Rules

To have a greater accuracy, the idea of the composite rule is to subdivide the interval \([a, b]\) into smaller intervals like \([x_{Qp}, x_{Qp+Q}]\), applying the quadrature formula in each of these smaller intervals and add up the results to obtain more accurate approximations.

\[
\int_a^b f(x)dx = \sum_{p=0}^{n_0-1} \int_{x_{Qp}}^{x_{Qp+Q}} f(x)dx
\]  

(2)

We define the Lagrange basis polynomials over the sub-interval \([x_{Qp}, x_{Qp+Q}]\).

\[
l_{Qp+j}(x) = \prod_{i\neq j=0}^{Q} \frac{x - x_{Qp+i}}{x_{Qp+j} - x_{Qp+i}} \quad l_j(x_i) = \delta_{ij} = \begin{cases} 0 : i \neq j \\ 1 : i = j \end{cases}
\]

(3)

The Lagrange Interpolating Polynomial and the integration can be derived

\[
\tilde{f}(x) = \sum_{j=0}^{Q} f(x_{Qp+j})l_{Qp+j}(x) \quad \int_{x_{Qp}}^{x_{Qp+Q}} \tilde{f}(x)dx = \sum_{j=0}^{Q} f(x_{Qp+j}) \int_{x_{Qp}}^{x_{Qp+Q}} l_{Qp+j}(x)dx
\]

(4)
The integration of Lagrange basis polynomials

$$\int_{x_{QP}}^{x_{QP+Q}} l_{QP+j}(x)dx = \frac{b-a}{n} (-1)^{Q-j} \int_{0}^{Q} \prod_{i \neq j=0}^{Q} (y-i)dy \quad (5)$$

We have the Lagrange Interpolating integration over $[x_{QP}, x_{QP+Q}]$

$$\int_{x_{QP}}^{x_{QP+Q}} \tilde{f}(x)dx = \frac{b-a}{n} \sum_{j=0}^{Q} W_j f(x_{QP+j}) \quad W_j = \frac{(-1)^{(Q-j)}}{j!(Q-j)!} \int_{0}^{Q} \prod_{i \neq j=0}^{Q} (y-i)dy \quad (6)$$

2.2. Error Analysis of the High Order Newton Cotes Formulas

It is shown in [13, 14] that for $f \in C^{Q+1}([x_{QP}, x_{QP+Q}])$, $x \in [x_{QP}, x_{QP+Q}]$, there exists $\eta(x) \in [x_{QP}, x_{QP+Q}]$, such that $f(x) - \tilde{f}(x) = \frac{1}{Q+1} f^{(Q+1)}(\eta(x)) \prod_{i=0}^{Q} (x-x_{QP+i})$

By integration, we have:

$$\int_{x_{QP}}^{x_{QP+Q}} f(x)dx - \int_{x_{QP}}^{x_{QP+Q}} \tilde{f}(x)dx = \frac{1}{Q+1} \int_{x_{QP}}^{x_{QP+Q}} f^{(Q+1)}(\eta(x)) \prod_{i=0}^{Q} (x-x_{QP+i})dx \quad (7)$$

For $Q$ even, [15] has shown in theorem 1.1 that for $f \in C^{Q+2}([x_{QP}, x_{QP+Q}])$, there exists $\eta_{p} \in [x_{QP}, x_{QP+Q}]$ such that

$$\frac{1}{Q+1} \int_{x_{QP}}^{x_{QP+Q}} f^{(Q+1)}(\eta(x)) \prod_{i=0}^{Q} (x-x_{QP+i})dx = \frac{f^{(Q+2)}(\eta_{p})}{Q+2!} \int_{x_{QP}}^{x_{QP+Q}} A^{*}(t)dt \quad (8)$$

with

$$A^{*}(t) = \int_{x_{QP}}^{t} A(x)dx \quad A(x) = \prod_{i=0}^{Q} (x-x_{QP+i}) \int_{x_{QP}}^{x_{QP+Q}} A^{*}(t)dt = h^{(Q+1)} \prod_{i=0}^{Q} (a-i) \quad (9)$$

The error analysis of the Newton-Cotes formulas of degree $Q$ in [16, 17] shows that the level of accuracy is greater when $Q$ is even. In fact, we have $O(h^{(Q+3)})$ for $Q$ even, against $O(h^{(Q+2)})$ for $Q$ odd. The estimation will be carried out with $Q$ even.

From (7), (8) and (9), we have the following integral on $[x_{QP}, x_{QP+Q}]$.

$$\int_{x_{QP}}^{x_{QP+Q}} f(x)dx = \int_{x_{QP}}^{x_{QP+Q}} \tilde{f}(x)dx + h^{Q+3} \frac{f^{(Q+2)}(\eta_{p})}{Q+2!} \int_{0}^{Q} \prod_{i=0}^{Q} (a-i)dady \quad (10)$$

**Proposition 1.1**

For $Q$ Even, $n = Qn_{0}$ integer, and $f \in C^{Q+2}([a, b])$, there exists $\eta \in [a, b]$ such that

$$\int_{a}^{b} f(x)dx = \frac{b-a}{n} \sum_{p=0}^{n-1} \sum_{j=0}^{Q} W_j f(x_{QP+j}) + h^{Q+2} \frac{f^{(Q+2)}(\eta)}{Q+2!} \int_{0}^{Q} \prod_{i=0}^{Q} (a-i)dady \quad (11)$$

With

$$W_j = \frac{(-1)^{(Q-j)}}{j!(Q-j)!} \int_{0}^{Q} \prod_{i \neq j=0}^{Q} (y-i)dy$$
Proof:  
From (5), we have the formula.

\[ \int_{a}^{b} f(x) \, dx = \sum_{p=0}^{n_0-1} \int_{x_{Qp}^p}^{r_{Qp}+Qh} f(x) \, dx \quad \text{From (10)} \]

\[ = \sum_{p=0}^{n_0-1} \int_{x_{Qp}^p}^{r_{Qp}+Q} \tilde{f}(x) \, dx + \frac{h^{Q+3}}{Q+2!} \left( \sum_{p=0}^{n_0-1} f^{(Q+2)}(\eta_p) \right) \int_{0}^{Q} \int_{0}^{Q} \prod_{i=0}^{Q} (a-i) \, dady \]

\[ = \sum_{p=0}^{n_0-1} \int_{x_{Qp}^p}^{r_{Qp}+Q} \tilde{f}(x) \, dx + h^{Q+3} \frac{n_0 f^{(Q+2)}(\eta)}{Q+2!} \int_{0}^{Q} \int_{0}^{Q} \prod_{i=0}^{Q} (a-i) \, dady \]

\[ = \frac{b-a}{n} \sum_{p=0}^{n_0-1} \sum_{j=0}^{Q} W_j f(x_{Qp+j}) + h^{Q+2} \frac{f^{(Q+2)}(\eta)}{(Q+2)!} \int_{0}^{Q} \int_{0}^{Q} \prod_{i=0}^{Q} (a-i) \, dady \quad \text{see (5)} \]

where \( n_0 = \frac{n}{Q} = \frac{b-a}{nQ} \) and \( \frac{1}{n_0} \sum_{p=0}^{n_0-1} f^{(Q+2)}(\eta_p) = f^{(Q+2)}(\eta) \) with \( \eta \in [a, b] \) (Intermediate Value Theorem).

2.3. Weights Computation  
Before using the formula in (11), we need to compute the weight \( \{W_j\}_{0 \leq j \leq Q} \) developed previously.

**Proposition 1.2**  
For \( Q \) Even and \( n = Qn_0 \) integer, \( j \in \{0, 1, 2, ..., Q\} \)

\[ W_j = \sum_{i=0}^{Q} C_i^j Q^{i+1} (-1)^{(Q-j)} (Q-j)! \]  

(12)

where \( (C_i^j)_{0 \leq i \leq Q} \) are coefficients of the polynomial functions.

Proof:  
For \( j \in \{0, 1, 2, ..., Q\} \), each polynomial function \( \prod_{i \neq j}^{Q} (y-i) \) can be developed as follows

\[ \prod_{i \neq j}^{Q} (y-i) = \sum_{i=0}^{Q} C_i^j y^j \]  

(13)

\[ W_j = \frac{(-1)^{(Q-j)}}{j!(Q-j)!} \int_{0}^{Q} \prod_{i \neq j}^{Q} (y-i) \, dy = \frac{(-1)^{(Q-j)}}{j!(Q-j)!} \sum_{i=0}^{Q} C_i^j Q^{i+1} \frac{1}{i+1} \]

\[ = \sum_{i=0}^{Q} C_i^j Q^{i+1} \frac{(-1)^{(Q-j)}}{i+1} j!(Q-j)! \]

The coefficients \( (C_i^j)_{0 \leq i \leq Q} \) were determined by resolving the equations (13) with Vandermonde matrix [18]. For \( Q = 12 \), the Lagrange polynomial is polynomial function of degree 12. The results of the weight \( \{W_j\}_{0 \leq j \leq Q} \) computations are summarized in the Table 1.
3. Numerical solution Analysis

The formula (11) will be used to compute the integral

$$
\int_{0}^{\infty} \left( \frac{(-z + k)\Gamma(k,z) + z^k e^{-z}}{k\Gamma(k)} \right)^{pM} dz
$$

(14)

In our case, \( f(x_{Qp+j}) = \left( \frac{(-x_{Qp+j} + k)\Gamma(k,x_{Qp+j}) + x_{Qp+j} e^{-x_{Qp+j}}}{k\Gamma(k)} \right)^{pM} \), \( n = Qn_0 \), \( 0 \leq p \leq n_0 - 1 \) and \( 0 \leq j \leq Q \). In order to compute the integral, the following parameter values are used: \( a = 0 \), \( b = 50 \), \( Q = 12 \), \( n_0 = 30000 \), \( 2500 \) and the weight \( \{W_j\}_{0 \leq j \leq Q} \) value in Table 1, second column. The upper bound \( (b = 50) \) yields good approximation results and is also efficient in term of computing time.

3.1. Compare Numerical and Analytical Solution for \( k=1 \)

For \( k = 1 \), the lifetime follows an exponential distribution and the analytical solution of the equation (1) is

$$
p = 1 - \frac{1}{R}
$$

(15)

**Proof:**

From (14) and (15), we have

$$
\int_{0}^{\infty} \left( \frac{(-z + k)\Gamma(k,z) + z^k e^{-z}}{k\Gamma(k)} \right)^{pM} dz = \int_{0}^{\infty} e^{-(1-\frac{1}{R})Mz} dz = \frac{1}{(1-\frac{1}{R})M} = \frac{1}{R(1-\frac{1}{R})\frac{1}{R}M}
$$

The numerical solution of the equation (1) was computed with reproduction number \( (R) \) as input. Newton-Cotes formula in (11) with \( Q = 12 \) was implemented to produce the numerical solution in Figure 2.

As shown in Figure 2 and Figure 3, the numerical and analytical solutions yield the same values at four decimal places. The Figure 4 shows an error between the two solutions with the maximum error \( 6 \times 10^{-5} \).
3.2. Infected size \((p)\) and Parameter shape \((k)\)

For \(k \neq 1\) and \(R = 2\), the equation (1) becomes

\[
\int_{0}^{\infty} \left( \frac{(-z + k)\Gamma(k, z) + z^k e^{-z}}{k\Gamma(k)} \right)^p M dz - \frac{k + 1}{4p(1-p)}M = 0 \tag{16}
\]

The shape parameter \(k\) is the input, The Newton-Cotes formula of 12 degrees was implemented to solve (16) numerically.

As shown in Figure 5, the infected size \((p)\) increases along with the shape parameter \((k)\). But the increase has two phases: an increasing rate phase and a decreasing rate phase. The shape of the infected size \((p)\) of Figure 5 might be explained by Figure 1.

When \(0 < k < 1\), the infected size \((p)\) is lower as the shape parameter \((k)\) is smaller. In fact, as shown in Figure 1 for \(0 < k < 1\), the instantaneous death rate is high and becomes higher as the shape parameter \((k)\) decreases. Therefore, the death process impact outweighs the infection process; and the infected size \((p)\) decreases when \(k\) lowers to 0, as shown in Figure 5.

When \(k > 1\), the level of infected size \((p)\) is higher and continues to increase, but at a decreasing rate, as shown in Figure 5. In fact, as shown in Figure 1 for \(k > 1\), the instantaneous death rate is lower and becomes lower as the shape parameter \((k)\) increases. Therefore, infected people live longer and contribute to the infection of healthy individual. We have fewer infected deaths, the new infected outweights the number of infected deaths. The infected size \((p)\) increases at a decreasing rate in Figure 5 as the infection process reaches its saturation state, which is defined by the reproduction number \((R = 2)\).
4. Contribution of the Extreme Value Theory
The Extreme Value Theory will consolidate the numerical solutions previously developed by providing an analytical formula of the equilibrium infected sized. However, the analytical solution is not valid for all the shape parameters $k$.

**Proposition 1.3**

Let $E_1(\infty), E_2(\infty), \ldots, E_n(\infty)$ be Independent and Identically Distributed (iid) positive random variables with cumulative distribution function $G(y) = \frac{1}{\mu} \int_0^y S(a)da$ and survival function $S(a) = \int_a^\infty f(y)dy$.

we have:

$$n \text{Min}(E_1(\infty), E_2(\infty), \ldots, E_n(\infty)) \text{ converges in distribution to an exponential distribution with parameter } \frac{1}{\mu}.$$ 

*Proof:*

$$P(n \text{Min}(E_1(\infty), E_2(\infty), \ldots, E_n(\infty)) \leq y) = P(\text{Min}(E_1(\infty), E_2(\infty), \ldots, E_n(\infty)) \leq \frac{y}{n})$$

$$= 1 - P(\text{Max}(E_1(\infty), E_2(\infty), \ldots, E_n(\infty)) > \frac{y}{n})$$

$$= 1 - \prod_{i=1}^{n} P(E_i(\infty) > \frac{y}{n}) = 1 - \prod_{i=1}^{n} (1 - G(\frac{y}{n}))$$

$$= 1 - (1 - G(\frac{y}{n}))^n$$

$$G(\frac{y}{n}) = \frac{1}{\mu} \int_0^{\frac{y}{n}} S(a)da = \frac{y}{\mu n} S(\eta) \text{ with } 0 \leq \eta \leq \frac{y}{n} \text{ (Mean Value Theorem for Integrals)}$$

$$P(n \text{Min}(E_1(\infty), E_2(\infty), \ldots, E_n(\infty)) \leq y) = 1 - (1 - G(\frac{y}{n}))^n = 1 - (1 - \frac{y}{\mu n} S(\eta))^n$$

When $n$ goes to $+\infty$, we have

$$\lim_{n \to +\infty} P(n \text{Min}(E_1(\infty), E_2(\infty), \ldots, E_n(\infty))) \leq y) = 1 - e^{\frac{1}{\mu} y} \quad (17)$$

We can conclude that

$$\lim_{n \to +\infty} E(n \text{Min}(E_1(\infty), E_2(\infty), \ldots, E_n(\infty))) = \mu \quad (18)$$

It was shown in [1] in Proposition 4.3.1 that

$$E(\text{Min}(I_1, I_2, \ldots, I_{M-n})) = \frac{1}{\beta n^{\frac{n}{M}}(M-n)}$$

The following equality was also stated in [1]

$$E(\text{Min}(E_1(\infty), E_2(\infty), \ldots, E_n(\infty))) = E(\text{Min}(I_1, I_2, \ldots, I_{M-n}))$$

From (18), we can develop the following expression

$$\mu = \lim_{n \to +\infty} E(n \text{Min}(E_1(\infty), E_2(\infty), \ldots, E_n(\infty)))$$

$$= \lim_{n \to +\infty} n E(\text{Min}(I_1, I_2, \ldots, I_{M-n}))$$
\[
\begin{align*}
&= \lim_{n \to +\infty} \frac{n}{\beta M (M - n)} \\
&= \frac{k + 1}{2R(1 - p)\lambda} \\
&= \frac{\mu k + 1}{2R(1 - p)k}
\end{align*}
\]

From the equation, the infected size \( p \) becomes
\[
p = 1 - \frac{k + 1}{k + \frac{1}{2R}}
\]

The analytical solution (19) is compared with the previous numerical solution. As shown in Figure (6), the analytical solution and the numerical solution produce the same value when \( k > 1 \). But, the analytical solution provides a solution limit \( p = 1 - \frac{1}{2R} \) as \( k \) goes to \( +\infty \).

![Figure 6. Numerical solution versus analytical solution of (16) for \( R = 2 \)](image)

When \( 0 \leq k \leq 1 \), on the contrary, the analytical solution does not match the numerical solution. In this case, the values of infected size \( p \) from analytical solution are negative and can not be the solution of the equation (16); whereas, the numerical solution gives better results for \( 0 \leq k \leq 1 \).

5. Conclusion
The effects of the Gamma distribution lifetime with parameter \( (k, \lambda) \) on the proportion \( (p) \) of the population infected at the equilibrium in a SIS model were our main interests in the paper. It results from the composite Newton-Cotes Quadrature formulas with \( Q = 12 \) degree, that the level of infected size \( p \) depends on the shape \( (k) \) parameter of the Gamma distribution. In fact, the infected size \( p \) increases along with the shape \( (k) \) parameter with an increasing rate and decreasing rate.

The shape of the infected size \( p \) solution can be explained by the instantaneous death rate characteristics impeded in the Gamma distribution hazard function in Figure 1. When \( 0 < k < 1 \), the instantaneous death rate is higher; we have the death process outweighs the infection process and resulting in low infection size \( p \). on the contrary, when \( k \geq 1 \), the
instantaneous death rate is lower, we have fewer infected death, and the new infected outweighs the number of infected deaths. The infected size (p) increases at a decreasing rate as the infection process reaches its saturation state, which is defined by the reproduction number (R). For k = 1, we have an exponential lifetime distribution, the numerical solution of the infected size yields the same values at four decimal places as the analytical solution (p = 1 − \frac{1}{R}).

The numerical solution was also compared with the analytical solution derived from the Extreme Value Theory. The results consolidate the numerical solutions of the infected size (p) when k ≥ 1 and provide a solution limit (p = 1 − \frac{1}{2R}) as k goes to +∞.

References
[1] Nzokem A H 2020 Stochastic and Renewal Methods Applied to Epidemic Models Ph.D. thesis York University, YorkSpace institutional repository
[2] Nzokem A H 2021 International Journal of Statistics and Probability 10 10–20
[3] Clancy D and Mendy S T 2011 Methodology and Computing in Applied Probability 13 603–618
[4] Allen L J 2008 An introduction to stochastic epidemic models Mathematical epidemiology (Springer) pp 81–130
[5] Ovaskainen O 2001 Journal of Applied Probability 38 898–907
[6] Wierman J C and Marchette D J 2004 Computational Statistics & Data Analysis 45 3–23
[7] Hernandez-Suarez C M, Castillo-Chavez C et al. 1997
[8] Næsell I 1999 Mathematical Biosciences 156 21–40
[9] Nzokem A and Madras N 2020 Bulletin of Mathematical Biology 82 9 1–16
[10] Nzokem A and Madras N 2021 Age-structured epidemic with adaptive vaccination strategy: Scalar-renewal equation approach Recent Developments in Mathematical, Statistical and Computational Sciences (Springer International Publishing) pp 591–599 ISBN 978-3-030-63591-6
[11] Yan P and Chowell G 2019 Quantitative methods for investigating infectious disease outbreaks vol 70 (Springer)
[12] Nzokem A H 2021 arXiv preprint arXiv:2104.07580 (Preprint 2104.07580)
[13] Shampine L F, Allen R C and Pruess S 1997 Fundamentals of Numerical Computing 2nd ed vol 1 (Wiley New York)
[14] Heath M T 2002 Scientific Computing: An Introductory Survey 2nd ed (New York : Academic Press)
[15] Hayes D and Rubin L 1970 The American Mathematical Monthly 77 1065–1072
[16] Omar A A S and Bashir M A 2015 International Journal of Scientific and Research Publications 5 1–6
[17] Krylov V I and Stroud A H 2006 Approximate calculation of integrals (Courier Corporation)
[18] Kalman D 1984 Mathematics Magazine 57 15–21