Proceedings Paper

Complementary Gamma Zero-Truncated Poisson Distribution and Its Application†

Ausaina Niyomdecha and Patchanok Srisuradetchai *

Department of Statistics, Faculty of Science and Technology, Thammasat University,
Pathum Thani, Thailand; ausaina.niy@dome.tu.ac.th
* Correspondence: patchanok@mathstat.sci.tu.ac.th
† Presented at the 1st International Online Conference on Mathematics and Applications; Available online: https://iocma2023.sciforum.net/.

Abstract: Numerous lifetime distributions have been developed to assist researchers in various fields. This paper proposes a new continuous three-parameter lifetime distribution called the complementary gamma zero-truncated Poisson distribution (CGZTP), which combines the distribution of the maximum of a series of independently identical gamma-distributed random variables with zero-truncated Poisson random variables. The proposed distribution’s properties, including proofs of the probability density function, cumulative distribution function, survival function, hazard function, and moments, are discussed. The unknown parameters are estimated using the maximum likelihood method, whose asymptotic properties are examined. In addition, Wald confidence intervals are constructed for the CGZTP parameters. Simulation studies are conducted to evaluate the efficacy of parameter estimation, and a real-world application demonstrates the application of the proposed distribution.

Keywords: compounding; gamma distribution; zero-truncated poisson distribution

1. Introduction

The gamma distribution is widely used in modeling lifetime data. However, the gamma distribution does not provide a reasonable parametric fit for modeling phenomena with non-monotone hazard rates, such as upside-down bathtub hazard rates. Some new distributions to model lifetime data have appeared in recent literature by compounding existing lifetime models with several discrete distributions. For instance, a distribution is obtained by assuming the minimum or maximum of continuous positive random variables. To accomplish this, Adamidis and Loukas [1] proposed an exponential-geometric (EG) distribution by compounding the geometric distribution and the exponential distribution. A complementary version of the EG distribution proposed by Louzada et al. [2], which would be applied to maximum lifetime data. The Weibull-geometric (WG) distribution with the minimum compounded function proposed by Barreto-Souza et al. [3] and its maximum version given by Tojeiro et al. [4]. Zakerzadeh and Mahmoudi [5] introduced a Lindley-geometric (LG) distribution, whereas Gui and Guo [6] introduced a complementary Lindley-geometric distribution. In addition, several new compoundings of Poisson distribution and some lifetime models have been introduced in closed forms, such as Kus [7], who proposed an exponential-Poisson (EP) distribution, Hemmati et al. [8], Lu and Chi [9], who proposed a Weibull-Poisson (WP), whose complementary version was given by Ismail [10]. Also, Alkarni and Oraby [11] introduced a Rayleigh-Poisson distribution, and Gui et al. [12] proposed a Lindley-Poisson distribution.

A novel distribution is established as the complementary gamma zero-truncated Poisson distribution (CGZTP). This paper is structured as follows: In Part 2, the distribution is mathematically derived, and in Section 3, its important properties are examined.
In Sections 4 and 5, the estimates of the parameters and the results of a simulation study are presented.

2. The Complementary Gamma Zero-Truncated Poisson Distribution

Let $X_1, X_2, \ldots, X_n$ be independent and identically distributed random variables from a gamma distribution which probability density function (pdf) given by

$$f(x; \alpha, \beta) = \beta^\alpha x^{\alpha-1} e^{-\beta x} / \Gamma(\alpha), \quad x > 0,$$

where $\alpha > 0$ is a shape parameter and $\beta > 0$ is a rate parameter, and $N$ is a random variable from zero-truncated Poisson distribution with probability mass function $P(N = n) = e^{-\lambda} \lambda^n / n! (1 - e^{-\lambda})$, $n = 1, 2, \ldots$ and $\lambda > 0$. Assuming that random variables $X$ and $N$ are independent, we define $Z = \max\{X_1, X_2, \ldots, X_n\}$.

Then, $g(z|\theta) = n \left[ F(z) \right]^{n-1} f(z)$, where $F(z) = 1 - \Gamma(\alpha, \beta z) / \Gamma(\alpha)$ and the marginal distribution for $Z$ is

$$g(z; \theta) = \frac{\lambda e^{-\lambda}}{1 - e^{-\lambda}} \left( \frac{\beta^\alpha z^{\alpha-1} e^{-\beta z}}{\Gamma(\alpha)} \right) e^{-\lambda \Gamma(\alpha, \beta z) / \Gamma(\alpha)}, \quad z > 0,$$

where $\theta = (\lambda, \alpha, \beta)$. The distribution of $Z$ will be referred to as CGZTP and plots of its pdf are displayed in Figure 1 for selected parameter values. For $\alpha = 1$, the CGZTP distribution reduces to the density of the complementary exponential-Poisson distribution introduced by Cancho et al.[13]. As $\lambda$ approaches to 0, the CGZTP distribution reduces to two-parameter gamma distribution.

![Figure 1](image_url)

**Figure 1.** Probability density functions of the CGZTP distribution for (a) $\alpha = 0.5$ and (b) $\alpha = 2$.

3. Properties of the Distribution

3.1. Cumulative Distribution Function, Quantile and Moment

The cumulative distribution function (cdf) of the CGZTP distribution is given by

$$G(z; \theta) = e^{-\lambda \left( \Gamma(\alpha, \beta z) / \Gamma(\alpha) \right) / \left( 1 - e^{-\lambda} \right)},$$

and the $r$th quantile is defined as the value $z$ such that

$$\Gamma(\alpha, \beta z) = -\Gamma(\alpha) \ln \left( r + \left( 1 - r \right) e^{-\lambda} \right) / \lambda.$$
In particular, the median is \( z \) such that \( \Gamma(\alpha, \beta z) = -\Gamma(\alpha)\ln\left(0.5 + 0.5e^{-1}\right)/\lambda \). Also, the moment generating function can be calculated from

\[
M_z(t) = \frac{\lambda \beta e^{\alpha \beta}}{\Gamma(\alpha)\left(1 - e^{-1}\right)} \int_0^{\infty} z^{\alpha - 1} e^{-\beta z - \frac{\Gamma(\alpha, \beta z)}{\Gamma(\alpha)}} dz.
\]

The \( k \) raw moments are given by

\[
E(Z^k) = \frac{\lambda \beta^k}{\Gamma(\alpha)\left(1 - e^{-1}\right)} \int_0^{\infty} z^{\alpha - 1 + k} e^{-\beta z - \frac{\Gamma(\alpha, \beta z)}{\Gamma(\alpha)}} dz, \ k \in \mathbb{N}.
\]

3.2. Survival Function and Hazard Function

Using Equations (1) and (2), the survival and hazard functions of the CGZTP distribution are given by

\[
S(z; \theta) = 1 - G(z; \theta) = \left(1 - e^{-\frac{\Gamma(\alpha, \beta z)}{\Gamma(\alpha)}} \right),
\]

and

\[
H(z; \theta) = \frac{g(z; \theta)}{S(z; \theta)} = \frac{\lambda \beta z^{\alpha - 1} e^{-\frac{\Gamma(\alpha, \beta z)}{\Gamma(\alpha)}}}{\Gamma(\alpha)\left(1 - e^{-1}\right)},
\]

respectively. If considering

\[
\eta(z) = \frac{g(z; \theta)}{S(z; \theta)} = -\left(1/z\right)\left[\alpha - 1 - \beta z + \lambda(\beta z)^{\alpha} e^{-\beta z}/\Gamma(\alpha)\right]
\]

and

\[
\eta'(z) = \frac{1}{\Gamma(\alpha)z^2}\left[(\alpha - 1)\Gamma(\alpha) + \lambda(\beta z)^{\alpha} (\beta z - \alpha + 1)e^{-\beta z}\right].
\]

For \( \alpha = 1 \), \( \eta'(z) > 0 \) for all \( z \), CGZTP distribution has an increasing hazard function that follows from Glaser [14]. Figure 2 illustrates some of the possible shapes of the hazard function for selected values of \( \theta \).

![Figure 2. Hazard functions of the CGZTP distribution for (a) \( \alpha = 2 \) and (b) \( \lambda = 2, \beta = 1 \).](image)

4. Parameter Estimation

4.1. Method of Maximum Likelihood

The log-likelihood function based on the observed random sample size of \( n \), \( w_{\text{obs}} = \left(z_1, z_2, \ldots, z_n\right) \) is the following:
\[ l(\theta; w_{ob}) = n \left( \log \lambda - \log \left( 1 - e^{-\lambda} \right) \right) + n\alpha \log \beta - n \log \Gamma(\alpha) + (\alpha - 1) \sum_{i=1}^{n} \log z_i \]

and the corresponding gradients are found to be

\[ \frac{\partial l(\theta; w_{ob})}{\partial \lambda} = n \left( \frac{1}{\lambda} - e^{-\lambda} \left( 1 - e^{-\lambda} \right)^{-1} \right) - \sum_{i=1}^{n} \Gamma(\alpha, \beta z_i) / \Gamma(\alpha), \tag{3} \]

\[ \frac{\partial l(\theta; w_{ob})}{\partial \alpha} = n \log \beta - n \psi_\alpha(\alpha) + \sum_{i=1}^{n} \log z_i \]

\[ -\lambda \sum_{i=1}^{n} \left[ G_{2,3}^{1,0} \left( \beta z_i ; 1,1 \right) + \Gamma(\alpha, \beta z_i) (\log(\beta z_i) - \psi_\alpha(\alpha)) \right] / \Gamma(\alpha), \tag{4} \]

\[ \frac{\partial l(\theta; w_{ob})}{\partial \beta} = \frac{n\alpha}{\beta} + \sum_{i=1}^{n} y_i - \frac{\lambda \beta^{-1}}{\Gamma(\alpha)} \sum_{i=1}^{n} z_i e^{-\beta y_i}, \tag{5} \]

where \( \psi(\alpha) \) is a digamma function and \( G_{p,q}^{m,n} \left( \beta z_i ; \beta \right) \) is Meijer-G-function. The equation (5) could be solved exactly for \( \lambda \), namely

\[ \hat{\lambda} = \frac{\Gamma(\hat{\lambda})}{\hat{\beta}^{b-1} \sum_{i=1}^{n} z_i e^{-\hat{\beta} y_i}} \left( \sum_{i=1}^{n} z_i - \frac{n \hat{\lambda}}{\hat{\beta}} \right) \]

conditional upon the value of \( \hat{\alpha} \) and \( \hat{\beta} \), where \( \hat{\lambda}, \hat{\alpha} \) and \( \hat{\beta} \) are maximum likelihood estimates for the parameter \( \lambda, \alpha \) and \( \beta \), respectively.

In the following theorem, some conditions are needed to be satisfied for the existence and uniqueness of the MLEs.

**Theorem 1.** Let \( l_i(\lambda; \alpha, \beta, w_{ob}) = \partial l(\theta; w_{ob}) / \partial \lambda \), If \( \alpha \) and \( \beta \) are known, then \( \hat{\lambda} \) is the uniquely exist root of \( l_i(\lambda; \alpha, \beta, w_{ob}) = 0 \) if \( \sum_{i=1}^{n} \Gamma(\alpha, \beta z_i) / \Gamma(\alpha) < n/2 \).

**Proof.** Because \( \lim_{\lambda \to 0} l_i(\lambda; \alpha, \beta, w_{ob}) > 0 \) as \( \sum_{i=1}^{n} \Gamma(\alpha, \beta z_i) / \Gamma(\alpha) < n/2 \) and \( \lim_{\lambda \to \infty} l_i(\lambda; \alpha, \beta, w_{ob}) < 0 \), there exist at least one solution of \( l_i(\lambda; \alpha, \beta, w_{ob}) = 0 \). Consider

\[ l'_i(\lambda; \alpha, \beta, w_{ob}) = -\frac{ne^4 (e^{-\lambda} + e^\lambda - (\lambda^2 + 2))}{e^4} \]

and use the fact that \( e^3 = 1 + \lambda + \frac{1}{2} \lambda^2 + \frac{1}{3!} \lambda^3 + \ldots \) and \( e^{-\lambda} = 1 - \lambda + \frac{1}{2} \lambda^2 - \frac{1}{3!} \lambda^3 + \ldots \), then \( e^{-\lambda} + e^{\lambda} = 2 + \lambda^2 + \frac{2}{4!} \lambda^4 + \ldots > \lambda^2 + 2 \). It follows that \( l'_i(\lambda; \alpha, \beta, w_{ob}) < 0 \) and \( l_i \) is strictly decreasing in \( \lambda \). Consequently, the root is proved to be unique. \( \Box \)
4.2. Variance-Covariance Matrix of the MLEs

The MLE of $\theta$ is approximately multivariate normal with a mean $\hat{\theta}$ and a variance-covariance matrix which is Fisher information matrix, i.e., $\hat{\theta} \sim N_3\left(\theta, \frac{1}{I(\hat{\theta})}\right)$ or $\hat{\theta} \sim N_3\left(\theta, I(\hat{\theta})^{-1}\right)$, where $I(\theta) = E\left[I(\theta)\right]$. $I(\theta)$ is the observed Fisher information matrix. By differentiating Equations (3)–(5), the elements of the observed Fisher information matrix are derived as follow:

$$I_{11} = n e^\alpha (e^\alpha + e^\beta - (\lambda^2 + 2))/e^\lambda,$$

$$I_{22} = n \psi^{(1)}(\alpha) + \lambda \sum_{i=1}^s \frac{1}{\Gamma(\alpha)} \left[ 2G_{2,4}^0 \left( \beta z_i, 1,1,0,0,\alpha \right) + 2 \left( \log(\beta z_i + \psi^{(0)}(\alpha) \right) \right]$$

$$I_{33} = \frac{n \alpha}{\beta^2} \frac{\lambda^{\alpha-2}}{\Gamma(\alpha)} \sum_{i=1}^s z_i^e e^{-\beta z_i} (\alpha - 1 - \beta z_i),$$

$$I_{12} = I_{21} = \frac{1}{\Gamma(\alpha)} \sum_{i=1}^s G_{2,3}^0 \left( \beta z_i, 1,1,0,0,\alpha \right) + \Gamma(\alpha, \beta z_i) (\log(\beta z_i + \psi^{(0)}(\alpha) - \psi^{(0)}(\alpha)^2 - \psi^{(0)}(\alpha) + \log^2(\beta z_i)).$$

$$I_{13} = I_{31} = -\lambda \beta^{\alpha-1} \left( \frac{1}{\Gamma(\alpha)} \right) \sum_{i=1}^s z_i^e e^{-\beta z_i},$$

$$I_{23} = I_{32} = -\frac{n}{\beta} - \lambda \sum_{i=1}^s e^{-\beta z_i} \left[ \frac{z_i^e \beta^{\alpha-1}}{\Gamma(\alpha)} \right] = -\frac{n}{\beta} - \lambda \sum_{i=1}^s e^{-\beta z_i} \left[ \frac{z_i^e \beta^{\alpha-1}}{\Gamma(\alpha)} \left( -\psi^{(0)}(\alpha) + \log(\beta) + \log(\beta) \right) \right].$$

To test the null hypothesis $H_0: \theta = \theta_0$, we can use Wald statistics: $J(\hat{\theta})^{1/2} (\hat{\theta} - \theta_0) \sim N_3(0, I_3)$ or $I(\hat{\theta})^{1/2} (\hat{\theta} - \theta_0) \sim N_3(0, I_3).$

The asymptotic distribution of the $i$th component of $\hat{\theta}$ is $\hat{\theta}_i \sim N(\theta_i, I^{\alpha})$ or $\hat{\theta}_i \sim N(\theta_i, l^a)$, where $l^a = \left[ I(\hat{\theta})^{-1} \right]_a$ and $I^{\alpha} = \left[ I(\hat{\theta})^{-1} \right]_a$. Then, the corresponding $(1 - \alpha)100\%$ Wald confidence intervals for $\theta_i$ are $\hat{\theta}_i \pm z_{1-a/2} \sqrt{l^a}$ or $\hat{\theta}_i \pm z_{1-a/2} \sqrt{l^a}.$

5. Simulation Study and Application

The study utilized 1000 simulated samples of 50, 100, and 1000. Table 1 shows the average MLEs of $\lambda, \alpha$, and $\beta$, and their MSEs when all parameters are unknown. As sample sizes increase, estimates become more accurate and MSE values decrease. Among three estimates, $\hat{\beta}$ tends to have smallest MSE.

Wald confidence intervals using observed Fisher information are constructed for all parameters of CGZTP. Monte Carlo simulations with 1000 repetitions help estimate the coverage probability (CP) and average length (AL) of the confidence intervals (CIs). All
results are presented in Table 2. It is found that when the sample size \( n \) increases, the CPs will be close to the nominal coverage probability, 0.95, and the ALs will decrease. If \( \lambda \) has a small value, i.e., \( \lambda = 0.5 \), \( n \) is required to be 1000 to achieve the nominal coverage probability. In most cases, coverage probabilities are close to 0.95.

Table 1. Mean estimates and mean-squared errors of \( \lambda, \alpha \), and \( \beta \).

<table>
<thead>
<tr>
<th>Distribution</th>
<th>( n )</th>
<th>( \hat{\lambda} )</th>
<th>( \hat{\alpha} )</th>
<th>( \hat{\beta} )</th>
<th>( \lambda )</th>
<th>( \hat{\alpha} )</th>
<th>( \hat{\beta} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>CGZTP(1,2,1)</td>
<td>50</td>
<td>1.7122</td>
<td>1.9225</td>
<td>1.0064</td>
<td>5.3198</td>
<td>0.4782</td>
<td>0.0519</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>1.6464</td>
<td>1.8901</td>
<td>0.9864</td>
<td>4.5325</td>
<td>0.3659</td>
<td>0.0272</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>1.0225</td>
<td>2.0003</td>
<td>0.9965</td>
<td>0.3992</td>
<td>0.0523</td>
<td>0.0025</td>
</tr>
<tr>
<td>CGZTP(3,1,0.5)</td>
<td>50</td>
<td>2.2571</td>
<td>1.4061</td>
<td>0.5503</td>
<td>4.4487</td>
<td>0.5402</td>
<td>0.0193</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>2.4622</td>
<td>1.2759</td>
<td>0.5294</td>
<td>3.4684</td>
<td>0.3293</td>
<td>0.0092</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>2.9382</td>
<td>1.0504</td>
<td>0.5049</td>
<td>0.8841</td>
<td>0.0597</td>
<td>0.0015</td>
</tr>
<tr>
<td>CGZTP(3,0.5,0.5)</td>
<td>50</td>
<td>2.3268</td>
<td>0.7150</td>
<td>0.5386</td>
<td>4.5848</td>
<td>0.1518</td>
<td>0.0165</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>2.5139</td>
<td>0.6553</td>
<td>0.5265</td>
<td>3.6790</td>
<td>0.0981</td>
<td>0.0086</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>2.7430</td>
<td>0.5561</td>
<td>0.5098</td>
<td>0.7855</td>
<td>0.0159</td>
<td>0.0010</td>
</tr>
</tbody>
</table>

For real-world use, the dataset is obtained from Proschan [15], and it is made up of 213 observations about how many times the air conditioning system on each of 13 Boeing 720 jet planes failed in a row. The CGZTP distribution was applied to the data, and \( \hat{\lambda} = 0.11, \hat{\alpha} = 0.84, \hat{\beta} = 0.01 \). The Kolmogorov-Smirnov (K-S) statistic is 0.05611 with \( p \)-values of 0.5230; therefore, the CGZTP distribution is useful for this data.

Table 2. Coverage probabilities and average lengths of Wald CIs.

<table>
<thead>
<tr>
<th>( \theta = (\lambda, \alpha, \beta) )</th>
<th>( n )</th>
<th>CP</th>
<th>AL</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \lambda = 0.5, \beta = 3 )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \alpha = 1 )</td>
<td>50</td>
<td>0.9130</td>
<td>1.4147</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.9090</td>
<td>1.0495</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>0.9530</td>
<td>0.3411</td>
</tr>
<tr>
<td>( \alpha = 2 )</td>
<td>50</td>
<td>0.9020</td>
<td>2.6400</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.8880</td>
<td>1.9795</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>0.9550</td>
<td>0.6626</td>
</tr>
<tr>
<td>( \lambda = 0.5 )</td>
<td>50</td>
<td>0.9800</td>
<td>6.0251</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.9610</td>
<td>4.5462</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>0.9520</td>
<td>1.3820</td>
</tr>
<tr>
<td>( \alpha = 1, \beta = 3 )</td>
<td>50</td>
<td>0.9840</td>
<td>6.2418</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.9580</td>
<td>5.2028</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>0.9560</td>
<td>1.7814</td>
</tr>
<tr>
<td>( \lambda = 1 )</td>
<td>50</td>
<td>0.9670</td>
<td>0.5587</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.9640</td>
<td>0.3832</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>0.9470</td>
<td>0.1144</td>
</tr>
<tr>
<td>( \lambda = 1, \alpha = 0.5 )</td>
<td>50</td>
<td>0.9710</td>
<td>1.0991</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.9630</td>
<td>0.7861</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>0.9530</td>
<td>0.2311</td>
</tr>
</tbody>
</table>

6. Conclusions

Gamma and zero-truncated Poisson are compounded to create the CGZTP distribution. The probability density function and hazard function plots showed this distribu-
tion’s flexibility. The MLEs and the corresponding variance-covariance matrix are mathematically derived. Furthermore, a simulation study was also conducted. Finally, the CGZTP model was applied to real data to demonstrate the distribution’s utility.

**Author Contributions:** All authors have read and agreed to the published version of the manuscript.

**Acknowledgments:** The authors would like to thank the editor and the reviewers for the valuable comments and suggestions to improve this paper. The research of Srisuradetchai is currently supported by the Thammasat University Research Unit in Theoretical and Computational Statistics.

**References**


**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.