

A New Extension of the Inverse Paralogistic Distribution using Gamma Generator with Application

Angelo E. Marasigan
Institute of Mathematical Sciences and Physics
University of the Philippines Los Baños
College, Laguna 4031 Philippines
aemarasigan@up.edu.ph

Date received: April 12, 2022

Revision accepted: March 28, 2023

Abstract

This study proposed a three-parameter model called the gamma inverse paralogistic (GiPL) distribution model. The probability density and cumulative distribution functions were presented together with the quantile function. Properties such as measures of reliability, the k^{th} raw moment and moment-generating function, partial moments, order statistics, log-likelihood functions for maximum likelihood estimations, Renyi entropy and the ordering of random variables were provided. To test the performance of the parameters, a simulation study was conducted. The simulation result was assessed using the mean, bias and root mean square errors. Finally, the data set on the number of COVID-19-infected individuals per age was used to apply the model and compared with various recently developed distribution models. Results showed the superiority of the GiPL distribution model over these models.

Keywords: gamma family, inverse paralogistic, real-data application, simulation study, statistical properties

1. Introduction

Many distributions were developed because of their wide range of applications to analyze various data sets in finance, actuarial science, hydrology, etc. These recent developments aim to build a distribution model that adequately and suitably captures the descriptive statistical properties of a data set. Furthermore, the model is capable to characterize its mean, median, mode, skewness, kurtosis and variability to control any of its shape, location and scale properties. Moreover, these recent developments add knowledge to the growing field of probability theory, particularly the use of probability models in data science and analytics, as these give more capability in the distribution modeling of data sets with sophisticated complexities.

Generalized distribution models have become evident in literature because of their capability to make new characterizations of the data set and where the models are shown to be suitable and efficient for that purpose. Some recently developed models are as follows: extended generalized Lindley distribution (Kantar *et al.*, 2018), generalized log-Moyal distribution (Bhati and Ravi, 2018), s generalized normal distribution (Cordeiro *et al.*, 2019), Marshall-Olkin Kappa distribution (Javed *et al.*, 2019), Marshall-Olkin length biased exponential distribution (ul Haq *et al.*, 2019), odd log-logistic Burr X distribution (Usman *et al.*, 2019), power Burr type X distribution (Usman and Ilyas, 2020), Kumaraswamy generalized Kappa distribution (Nawaz *et al.*, 2020), Marshall-Olkin extended inverted Kumaraswamy distribution (Usman and ul Haq, 2020), Marshall-Olkin power Lomax distribution (ul Haq *et al.*, 2020) and Marshall-Olkin inverted Nadarajah-Haghighi distribution (Raffiq *et al.*, 2020).

In this study, a three-parameter distribution called the gamma inverse paralogistic (GIPL) model was developed. The formulation was based on generating a particular member of a gamma family using the inverse paralogistic distribution as the base model. The two-parameter distribution model called the inverse paralogistic model has the following respective probability density function (Equation 1) and cumulative distribution function (Equation 2):

$$g(x; \tau, \lambda) = \frac{\tau^2 (\lambda x) \tau^2}{x [1 + (\lambda x) \tau]^{\tau+1}} \tag{1}$$

$$G(x; \tau, \lambda) = \left[\frac{(\lambda x) \tau}{1 + (\lambda x) \tau} \right]^\tau \tag{2}$$

where $x > 0$ and $\tau, \lambda > 0$ are the parameters. In the inverse paralogistic model, λ serves as the shape parameter and τ is the scale parameter. The figure below shows the plot of Equation 1 of this model.

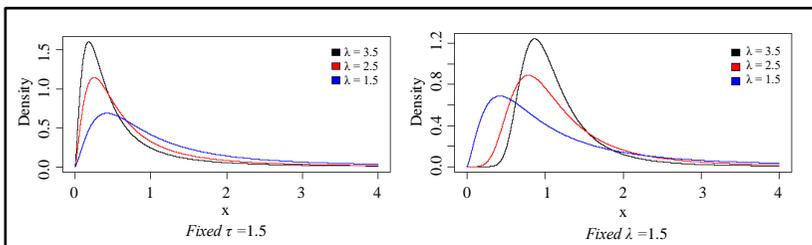


Figure 1. Probability density function of the inverse paralogistic distribution

The gamma family is based on the following structure: let $G(x)$ be the cumulative distribution function of a random variable X with probability density function $g(x)$. Then, Zografos and Balakrishnan (2009) formulated the gamma family distribution, which has the following probability density function and cumulative density function defined by Equations 3 and 4, respectively.

$$f(x; \alpha) = \frac{1}{\Gamma(\alpha)} \{-\ln[1 - G(x)]\}^{\alpha-1} g(x) \tag{3}$$

$$F(x; \alpha) = \frac{\gamma_z\{\alpha, -\ln[1 - G(x)]\}}{\Gamma(\alpha)} \tag{4}$$

where $\alpha > 0$ is a parameter, $\Gamma(\cdot)$ is the gamma function, and for each $z > 0$,

$$\gamma_-(\alpha, z) = \int_0^z t^{\alpha-1} e^{-t} dt$$

is the lower incomplete gamma function. As an illustration, say G and g are the cumulative distribution and probability density functions, respectively, of a standard normal random variable. Then, Figure 2 shows Equation 3 with a normal distribution base model.

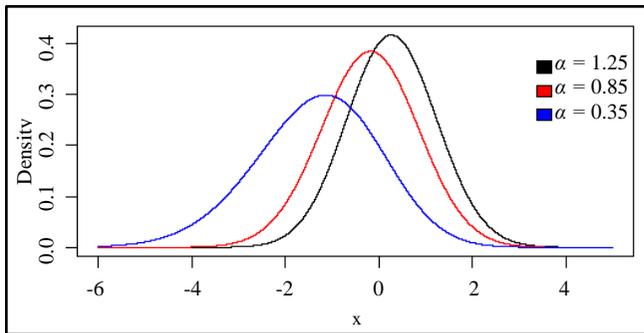


Figure 2. Probability density function of a gamma family with normal distribution base model

Some of the models available in the literature with properties developed and with illustration for application are the following: gamma generalized Pareto distribution (De Andrade *et al.*, 2017) and gamma Burr XII distribution (Guerra *et al.*, 2017).

The development of the model was motivated by the limited available study regarding the extension of the inverse paralogistic distribution model, which

is classically used to model data sets that are heavy-tailed. Commonly, data sets in hydrology, finance and actuarial science depict such behavior. In this model, the additional parameter makes it more flexible to handle a data set with sophisticated behavior regarding its shape and scale.

2. Methodology

2.1 Model Description

A new model called the GiPL distribution is a three-parameter model based on the inverse paralogistic model with a family generator. By substituting Equation 1 and Equation 2 into Equation 3, the probability density function is defined by

$$f(x; \alpha, \tau, \lambda) = \frac{1}{\Gamma(\alpha)} \left\{ -\ln \left[1 - \left[\frac{(\lambda x)^\tau}{1 + (\lambda x)^\tau} \right]^\tau \right] \right\}^{\alpha-1} \frac{\tau^2 (\lambda x)^{\tau^2}}{x [1 + (\lambda x)^\tau]^{\tau+1}} \quad (5)$$

where $x > 0$. As for the cumulative distribution function, we substitute Equation 2 into Equation 4 to get

$$F(x; \alpha, \tau, \lambda) = \frac{1}{\Gamma(\alpha)} \gamma_{-} \left\{ \alpha, -\ln \left[1 - \left[\frac{(\lambda x)^\tau}{1 + (\lambda x)^\tau} \right]^\tau \right] \right\} \quad (6)$$

We denote a random variable X following a GiPL distribution with parameters α , τ , and λ as $X \sim \text{GiPL}(\alpha, \tau, \lambda)$. The new model has more control over the shape of the distribution due to α and λ , and provides characteristics about the spread of the distribution due to the scale parameter τ . Figure 3 shows the graph of the probability density function with varying values of the parameters.

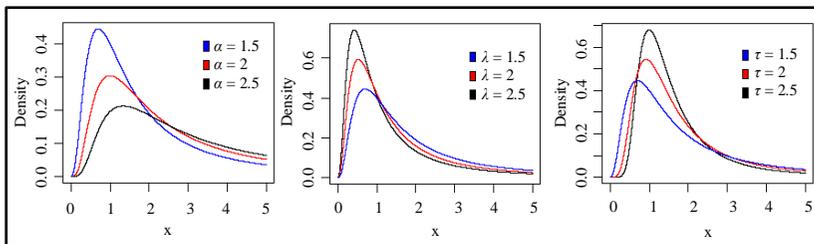


Figure 3. Plots of the probability density function of $\text{GiPL}(\alpha, \tau, \lambda)$ with varying parameters

The following result shows the quantile function for a GiPL(α, τ, λ) distribution. Specifically, the quantile function Q for X at a level p , given in Equation 7, is formulated as follows: by letting $p: F(\bar{x}, \alpha, \tau, \lambda)$, for a fixed $\bar{x} > 0$ and using Equation 6, we get

$$p = \frac{I}{\Gamma(\alpha)} \gamma_- \left\{ \alpha, -\ln \left[I - \left[\frac{(\lambda \bar{x})^\tau}{I + (\lambda \bar{x})^\tau} \right]^\tau \right] \right\}$$

Solving for \bar{x} leads to

$$Q(p) = \frac{I}{\lambda} \left\{ \left[I - \exp\{-\gamma_-^{-1}(\alpha, p\Gamma(\alpha))\} \right]^{-1/\tau} - I \right\}^{-1/\tau} \tag{7}$$

where γ_-^{-1} is the inverse of the lower incomplete gamma function.

2.2 Data Preparation

In this study, GiPL(α, τ, λ) was fitted to a data set. The data set used is on the number of COVID-19-infected individuals per age from April 2020 to March 2022 in the National Capital Region (NCR) of the Philippines – a highly populated region in the country. The data was retrieved from the Philippine COVID-19 Data Dashboard made by the University of the Philippines Los Baños Biomathematics Initiative (Biomathematics Research Cluster, 2022).

In summary, Table 1 presents some descriptive statistical measures for the data.

Table 1. Descriptive statistical measures using the data set

Data	Mean	Median	Mode	Variance	Skewness	Kurtosis
COVID-19	38.536	38	32	308.5909	0.0340	3.3336

As observed, mode < median < mean for the given data set. Additionally, the skewness was positive. Hence, the distribution of the data set was positively skewed. Furthermore, the variance was relatively large which means that data points tended to be far from the mean. Lastly, because the kurtosis was more than one, the distribution of the data set was too peaked compared with the normal curve.

2.3 Development of Properties of $\text{GiPL}(\alpha, \tau, \lambda)$

The study provided various mathematical and statistical properties of a random variable $X \sim \text{GiPL}(\alpha, \tau, \lambda)$. Specifically, some measures of reliability such as the survival, hazard and cumulative hazard function were derived. These were followed by the derivation of the k^{th} raw moment and moment-generating function of X . Then, the partial moments were derived, and the order statistics were formulated and presented. For the estimation of parameters using maximum likelihood estimation, the log-likelihood function was provided with the score vector: the vector of partial derivatives with respect to the parameters. The information matrix was also provided here. The stochastic, likelihood-ratio and hazard orderings were established as well. Finally, given that the Renyi entropy has no closed form, an approximation was formulated.

2.4 Parameter Estimation Procedure

To apply the $\text{GiPL}(\alpha, \tau, \lambda)$ model, calibration of the model to the given data set was illustrated. The method used to estimate the parameters is a popular technique known as the maximum likelihood estimation, which uses the log-likelihood function of the distribution model and proceeds with the maximization technique. It is uncommon for complex maximization problems to have analytic solutions. This study employed numerical techniques using R software (R Core Team, 2023) to estimate the parameters along with statistical measures to verify the accuracy and adequacy of the model. The performance of this estimation procedure was backed by the conduct of a simulation study supported by various measures such as the mean, bias and root mean square errors (RMSE). Lastly, the data set on the number of COVID-19-infected individuals was used to illustrate the model.

3. Results and Discussion

3.1 Mathematical Properties

3.1.1 Reliability Analysis

Let $X \sim \text{GiPL}(\alpha, \tau, \lambda)$. We provide here the survival, hazard and cumulative hazard functions of X . The survival function, denoted by S , is given by

$$\begin{aligned}
 S(x; \alpha, \tau, \lambda) &= -I F(x; \alpha, \tau, \lambda) \\
 &= \frac{I}{\Gamma(\alpha)} \gamma_+ \left\{ \alpha, -\ln \left[I - \left[\frac{(\lambda x)^\tau}{I + (\lambda x)^\tau} \right]^\tau \right] \right\}
 \end{aligned}
 \tag{8}$$

where γ_+ is the upper incomplete gamma function:

$$\gamma_+(\alpha, z) = \int_z^{+\infty} t^{\alpha-1} e^{-t} dt$$

Figure 4 shows the graph of the survival function with varying values of the parameters.

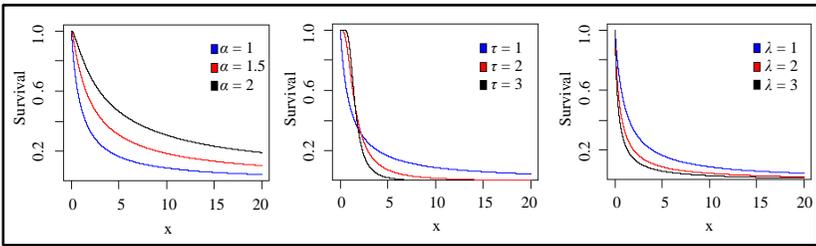


Figure 4. Plot of the survival function of $GiPL(\alpha, \tau, \lambda)$ with varying parameters

The hazard function, denoted by h , is the ratio of the probability density and the survival functions. It is given by

$$h(x; \alpha, \tau, \lambda) = \frac{\left\{ -\ln \left[I - \left[\frac{(\lambda x)^\tau}{I + (\lambda x)^\tau} \right]^\tau \right] \right\}^{\alpha-1} \frac{\tau^2 (\lambda x)^{\tau^2}}{x [I + (\lambda x)^\tau]^{\tau+1}}}{\gamma_+ \left\{ \alpha, -\ln \left[I - \left[\frac{(\lambda x)^\tau}{I + (\lambda x)^\tau} \right]^\tau \right] \right\}}
 \tag{9}$$

Figure 5 shows the graph of the hazard function with varying values of the parameters.

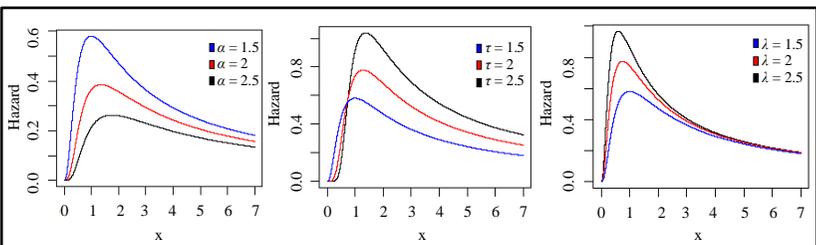


Figure 5. Plot of the hazard function of $GiPL(\alpha, \tau, \lambda)$ with varying parameters (other parameters are fixed at 1.5.)

Finally, the cumulative hazard function, denoted by ch , is the negative logarithm of the survival function. It is given by

$$ch(x; \alpha, \tau, \lambda) = \ln[\Gamma(\alpha)] - \ln \left\{ \gamma_+ \left\{ \alpha, -\ln \left[1 - \left[\frac{(\lambda x)^\tau}{1 + (\lambda x)^\tau} \right]^\tau \right] \right\} \right\} \quad (10)$$

3.1.2 Moments

The k^{th} raw moment of $X \sim \text{GiPL}(\alpha, \tau, \lambda)$ is summarized in the following theorem:

Theorem 1 (k^{th} raw moment). Let μ_k be the k^{th} raw moment of $X \sim \text{GiPL}(\alpha, \tau, \lambda)$. Then

$$\mu_k = \frac{1}{\lambda^k} \sum_{r=0}^{+\infty} \sum_{s=0}^{+\infty} \binom{z_1}{r} \binom{z_2}{s} (-1)^{z_1+z_2-(r+s)-r/\tau} \frac{1}{(I+s)^\alpha} \quad (11)$$

where $z_1 = -\frac{k}{\tau}$ and $z_2 = -\frac{r}{\tau}$.

Proof.

From Equation 5, it follows that

$$\begin{aligned} \mu_k &= \int_0^{+\infty} x^k f(x; \alpha, \tau, \lambda) dx \\ &= \int_0^{+\infty} x^k \frac{1}{\Gamma(\alpha)} \left\{ -\ln \left[1 - \left[\frac{(\lambda x)^\tau}{1 + (\lambda x)^\tau} \right]^\tau \right] \right\}^{\alpha-1} \frac{\tau^2 (\lambda x)^{\tau^2}}{x [1 + (\lambda x)^\tau]^{\tau+1}} dx \quad (12) \end{aligned}$$

Let $= -\ln \left[1 - \left[\frac{(\lambda x)^\tau}{1 + (\lambda x)^\tau} \right]^\tau \right]$. Then Equation 12 implies that

$$\begin{aligned} \mu_k &= \int_0^{+\infty} \frac{1}{\lambda^k} \left[[1 - e^{-u}]^{\frac{1}{\tau}} - 1 \right]^{-k/\tau} \frac{1}{\Gamma(\alpha)} u^{\alpha-1} e^{-u} du \\ &= \frac{1}{\lambda^k \Gamma(\alpha)} \int_0^{+\infty} \sum_{r=0}^{+\infty} (1 - e^{-u})^{-\frac{r}{\tau}} (-1)^{z_1-r} \binom{z_1}{r} u^{\alpha-1} e^{-u} du \quad (13) \end{aligned}$$

$$= \frac{1}{\lambda^k \Gamma(\alpha)} \int_0^{+\infty} \sum_{r=0}^{+\infty} \sum_{s=0}^{+\infty} e^{-us} (-1)^{z_2-s} \binom{z_2}{s} (-1)^{z_1-r-r/\tau} \binom{z_1}{r} u^{\alpha-1} e^{-u} du \quad (14)$$

$$= \frac{1}{\lambda^k \Gamma(\alpha)} \sum_{r=0}^{+\infty} \sum_{s=0}^{+\infty} (-1)^{z_2-s} \binom{z_2}{s} (-1)^{z_1-r-r/\tau} \binom{z_1}{r} \int_0^{+\infty} u^{\alpha-1} e^{-u(I+s)} du \quad (15)$$

where $z_1 = -\frac{k}{\tau}$ and $z_2 = -\frac{r}{\tau}$. Equations 13 and 14 follow from Newton's Generalization of the Binomial Theorem.

Now, let $y = u(1 + s)$. Then (15) implies (11), as desired. ■

The next theorem gives the moment-generating function $M(t)$ of $X \sim \text{GiPL}(\alpha, \tau, \lambda)$ and the final expression uses the relationship (Equation 16).

$$\begin{aligned}
 M(t) &= \int_0^{+\infty} e^{tx} f(x; \alpha, \tau, \lambda) dx \\
 &= \int_0^{+\infty} \sum_{k=0}^{+\infty} \frac{(tx)^k}{k!} f(x; \alpha, \tau, \lambda) dx \\
 &= \sum_{k=0}^{+\infty} \frac{t^k}{k!} \mu_k
 \end{aligned} \tag{16}$$

Theorem 2 (Moment-generating function). Let $X \sim \text{GiPL}(\alpha, \tau, \lambda)$ and $M(t)$, $t > 0$, denote the moment-generating function of X . Then,

$$M(t) = \sum_{k=0}^{+\infty} \sum_{r=0}^{+\infty} \sum_{s=0}^{+\infty} \frac{t^k}{\lambda^k k!} \binom{z_1}{r} \binom{z_2}{s} (-1)^{z_1+z_2-(r+s)-r/\tau} \frac{1}{(1+s)^\alpha} \tag{17}$$

where $z_1 = -\frac{k}{\tau}$ and $z_2 = -\frac{r}{\tau}$.

Proof.

The proof is similar to the proof of Theorem 1. ■

3.1.3 Partial Moments

In general, if X is a random variable with continuous probability density function f over its domain and with parameter vector θ , then the lower and upper partial moments with respect to v are, respectively, defined by solving the following integrals:

$$\int_0^v (v-x)^k f(x; \theta) dx \quad \text{and} \quad \int_v^{+\infty} (x-v)^k f(x; \theta) dx.$$

The k^{th} lower and upper partial moments of $X \sim \text{GiPL}(\alpha, \tau, \lambda)$ are summarized in the following theorem:

Theorem 3 (Lower and upper partial moments). Let $X \sim \text{GiPL}(\alpha, \tau, \lambda)$ and denote $\mu_k^-(v)$ and $\mu_k^+(v)$, respectively, the lower and upper partial moments of X with respect to v . Then,

$$\mu_k^-(v) = \sum_{j=0}^k \sum_{r=0}^{+\infty} \sum_{s=0}^{+\infty} \binom{k}{j} \binom{y_1}{r} \binom{y_2}{s} v^j (-1)^q \frac{v - [\alpha, \bar{u}(I+s)]}{(I+s)^\alpha \lambda^{k-j} \Gamma[\alpha]} \tag{18}$$

where $\bar{u} = -\ln \left[I - \left[\frac{(\lambda v)^\tau}{I + (\lambda v)^\tau} \right]^\tau \right]$, $y_1 = -\frac{k-j}{\tau}$, $y_2 = -\frac{r}{\tau}$ and $q = k + y_1 + 2y_2 - (j+r+s)$, and

$$\mu_k^+(v) = \sum_{j=0}^k \sum_{r=0}^{+\infty} \sum_{s=0}^{+\infty} \binom{k}{j} \binom{t_1}{r} \binom{t_2}{s} v^{k-j} (-1)^p \frac{v + [\alpha, \bar{u}(I+s)]}{(I+s)^\alpha \lambda^j \Gamma[\alpha]} \tag{19}$$

where $t_1 = -\frac{j}{\tau}$, $t_2 = -\frac{r}{\tau}$ and $p = k + t_1 + 2t_2 - (j+r+s)$.

Proof.

Following the definition of the lower incomplete partial moment and using Equation 5, we get

$$\begin{aligned} \mu_k^-(v) &= \int_0^v (v-x)^k f(x; \alpha, \tau, \lambda) dx \\ &= \int_0^v (v-x)^k \frac{1}{\Gamma(\alpha)} \left\{ -\ln \left[I - \left[\frac{(\lambda x)^\tau}{I + (\lambda x)^\tau} \right]^\tau \right] \right\}^{\alpha-1} \frac{\tau^2 (\lambda x)^\tau}{x [I + (\lambda x)^\tau]^{\tau+1}} dx \\ &= \sum_{j=0}^k \binom{k}{j} v^j (-1)^{k-j} \int_0^v \frac{1}{\Gamma(\alpha)} \left\{ -\ln \left[I - \left[\frac{(\lambda x)^\tau}{I + (\lambda x)^\tau} \right]^\tau \right] \right\} \times \frac{\tau^2 (\lambda x)^\tau}{x^{I+j-k} [I + (\lambda x)^\tau]^{\tau+1}} dx \end{aligned} \tag{20}$$

Equation 20 follows from the binomial expansion of $(v-x)^k$.

Let $u = -\ln \left[I - \left[\frac{(\lambda x)^\tau}{I + (\lambda x)^\tau} \right]^\tau \right]$. Then Equation 20 implies that

$$\mu_k^-(v) = \sum_{j=0}^k \binom{k}{j} v^j (-1)^{k-j} \int_0^{\bar{u}} \frac{1}{\lambda^{k-j}} \left[(I - e^{-u})^{\frac{1}{\tau}} - I \right]^{-\frac{k-j}{\tau}} \times \frac{1}{\Gamma(\alpha)} u^{\alpha-1} e^{-u} du \tag{21}$$

where $\bar{u} = -\ln \left[I - \left[\frac{(\lambda v)^\tau}{I + (\lambda v)^\tau} \right]^\tau \right]$. By Newton's generalization of the binomial theorem, we get from Equation 21 that

$$\begin{aligned} \mu_k(v) &= \sum_{j=0}^j \binom{k}{j} v^j (-1)^{k-j} \int_0^{\bar{u}} \frac{I}{\lambda^{k-j}} \sum_{r=0}^{+\infty} \sum_{s=0}^{+\infty} \binom{y_1}{r} \binom{y_2}{s} (-1)^{y_1+2y_2-(r+s)} \times \frac{I}{\Gamma(\alpha)} u^{\alpha-1} e^{-u} du \\ &= \sum_{j=0}^j \sum_{r=0}^{+\infty} \sum_{s=0}^{+\infty} \binom{k}{j} \binom{y_1}{r} \binom{y_2}{s} v^j (-1)^q \frac{I}{\lambda^{k-j} \Gamma(\alpha)} \int_0^{\bar{u}} u^{\alpha-1} e^{-u} du \end{aligned} \tag{22}$$

where $y_1 = -\frac{k-j}{\tau}$, $y_2 = -\frac{r}{\tau}$, and $q = k + y_1 + 2y_2 - (j + r + s)$. Using the notation for the lower incomplete gamma function, Equation 22 implies Equation 18, as desired.

Now the upper incomplete partial moment is defined by

$$\begin{aligned} \mu_k^-(v) &= \int_v^{+\infty} (x-v)^k f(x; \alpha, \tau, \lambda) dx \\ &= \int_v^{+\infty} (x-v)^k \frac{I}{\Gamma(\alpha)} \left\{ -\ln \left[I - \left[\frac{(\lambda x)^\tau}{I + (\lambda x)^\tau} \right]^\tau \right] \right\}^{\alpha-1} \frac{\tau^2 (\lambda x)^{\tau^2}}{x [I + (\lambda x)^\tau]^{\tau+1}} dx. \end{aligned}$$

The upper incomplete partial moment given in the theorem is developed similarly as the lower incomplete partial moment (Equation 18). ■

3.1.4 Order Statistics

As a preliminary, let $Y_{(j)}$, $j = 1, 2, \dots, n$ be the order statistics of the random samples Y_i , $i = 1, 2, \dots, n$, taken from the cumulative distribution function model F . The cumulative distribution function $h_{j,n}(y)$ of the j^{th} order statistic $Y_{(j)}$ is given by

$$h_{j,n}(y) = \sum_{k=j}^n \binom{n}{k} [F(y)]^k [I - F(y)]^{n-k}. \tag{23}$$

Reindexing the summation from $k = j$ to $k = 0$ makes the right-hand side of Equation 23 becomes

$$h_{j,n}(y) = \sum_{k=0}^{n-j} \binom{n}{k+j} [F(y)]^{k+j} [I - F(y)]^{n-k-j}. \tag{24}$$

Observe that

$$\begin{aligned} \binom{n}{k+j} &= \frac{n!}{(n-k-j)!(k+j)!} \\ &= \frac{(n-j)!}{(n-k-j)!k!} \frac{P_j^n}{P_{k+j}^n} \\ &= \binom{n-j}{k} \frac{P_j^n}{P_{k+j}^n} \end{aligned} \tag{25}$$

where $P_b^a = \frac{a!}{(a-b)!}$, for each integer $0 \leq b \leq a$. Substituting Equation 25 to Equation 24 and doing some algebra result to

$$h_{j,n}(y) = \frac{P_j^n}{P_{k+j}^n} \left[\frac{F(y)}{1-F(y)} \right]^j = \frac{P_j^n}{P_{k+j}^n} \left[\frac{F(y)}{S(y)} \right]^j. \tag{26}$$

To derive the cumulative distribution function of the j^{th} order statistics of a random variable with a desired distribution, Equation 26 is used. Now, let X_i , $i=1, 2, \dots, n$, be random samples from a GiPL(α, τ, λ) distribution, where $X_{(j)}$, $j=1, 2, \dots, n$ are the order statistics. That is, $X_{(k)} < X_{(l)}$ for each $1 \leq k < l \leq n$. Using Equation 26 with the distribution model of $X \sim \text{GiPL}(\alpha, \tau, \lambda)$ given in Equation 6, we get

$$h_{j,n}(y) = \frac{P_j^n}{P_{k+j}^n} \left[\frac{y - \left\{ \alpha, -\ln \left[1 - \left[\frac{(\lambda x)^\tau}{1 + (\lambda x)^\tau} \right]^\tau \right] \right\}}{\gamma + \left\{ \alpha, -\ln \left[1 - \left[\frac{(\lambda x)^\tau}{1 + (\lambda x)^\tau} \right]^\tau \right] \right\}} \right]^j \tag{27}$$

the cumulative distribution function of the j^{th} order statistic of $X_{(j)}$.

3.1.5 Log-likelihood Function

For a random variable with density $f(x)$ and parameter vector θ , the log-likelihood function is $\ell(\theta; x)$. Using Equation 5 and letting $\theta = (\alpha, \tau, \lambda)^T$, we get its log-likelihood function given by

$$\ell(\theta; x) = -\ln \Gamma(\alpha) + (\alpha - 1) \ln \left\{ -\ln \left[1 - \left[\frac{(\lambda x)^\tau}{1 + (\lambda x)^\tau} \right]^\tau \right] \right\} + 2 \ln \tau + \tau^2 (\ln \lambda - \ln x) - \ln x - (\tau + 1) \ln [1 + (\lambda x)^\tau]. \tag{28}$$

A method to estimate the parameters of a distribution model is by using maximum likelihood estimation (MLE), which uses the log-likelihood function of a random variable. Maximizing the log-likelihood function with respect to the parameters of the distribution model establishes MLE. Given a

data set $\{x_1, x_2, \dots, x_n\}$, the maximum likelihood estimator $\hat{\theta}$ is generated by maximizing $\ell_n(\theta; x) = \sum_{i=1}^n \ell(\theta; x_i)$.

Now, using Equation 28, we get the score vector ∇_{θ} , which is defined by

$$\nabla_{\theta} = \left(\frac{\partial \ell(\theta; x)}{\partial \alpha}, \frac{\partial \ell(\theta; x)}{\partial \tau}, \frac{\partial \ell(\theta; x)}{\partial \lambda} \right)^T.$$

Alternatively, the estimators by MLE can be derived by solving $\nabla_{\theta} = \vec{0}$. Taking the partial derivatives, we get

$$\begin{aligned} \frac{\partial \ell(\theta; x)}{\partial \alpha} &= \frac{-d(\alpha)}{\Gamma(\alpha)} + \ln \left\{ -\ln \left[I - \left(\frac{(\lambda x)^{\tau}}{I + (\lambda x)^{\tau}} \right)^{\tau} \right] \right\} \\ \frac{\partial \ell(\theta; x)}{\partial \tau} &= \left\{ -\ln \left[I - \left(\frac{(\lambda x)^{\tau}}{I + (\lambda x)^{\tau}} \right)^{\tau} \right] \right\}^{-1} \left[I - \left(\frac{(\lambda x)^{\tau}}{I + (\lambda x)^{\tau}} \right)^{\tau} \right]^{-1} \left(\frac{(\lambda x)^{\tau}}{I + (\lambda x)^{\tau}} \right)^{\tau} \\ &\quad \times \ln \left(\frac{(\lambda x)^{\tau}}{I + (\lambda x)^{\tau}} \right) \frac{(\lambda x)^{\tau} \ln(\lambda x)}{[I + (\lambda x)^{\tau}]^2} + \frac{2}{\tau} + 2\tau(\ln \lambda + \ln x) - (\tau + 1) \\ &\quad \times [I + (\lambda x)^{\tau}]^{-1} (\lambda x)^{\tau} \ln(\lambda x)^{\tau} - \ln [I + (\lambda x)^{\tau}] \\ \frac{\partial \ell(\theta; x)}{\partial \tau} &= (\alpha - I) \left\{ -\ln \left[I - \left(\frac{(\lambda x)^{\tau}}{I + (\lambda x)^{\tau}} \right)^{\tau} \right] \right\}^{-1} \left[I - \left(\frac{(\lambda x)^{\tau}}{I + (\lambda x)^{\tau}} \right)^{\tau} \right]^{-1} \\ &\quad \times \tau \left(\frac{(\lambda x)^{\tau}}{I + (\lambda x)^{\tau}} \right)^{\tau-1} \frac{\tau x (\lambda x)^{\tau-1}}{[I + (\lambda x)^{\tau}]^2} + \frac{\tau^2}{\lambda} - (\tau + I) [I + (\lambda x)^{\tau}]^{-1} \tau x (\lambda x)^{\tau-1} \end{aligned}$$

The estimates are obtained by simultaneously solving the equations

$$\frac{\partial \ell(\theta; x)}{\partial \alpha} = 0, \quad \frac{\partial \ell(\theta; x)}{\partial \tau} = 0, \quad \frac{\partial \ell(\theta; x)}{\partial \lambda} = 0$$

The closed-form solution is not available; hence, a numerical method may be employed (i.e. Simulated Annealing, Newton-Quasi, Broyden-Fletcher-Goldfarb-Shanno method, Nelder-Mead, etc.) to approximate the estimates of the parameters.

Now, the information matrix, denoted by $J(\theta)$, is defined by

$$J(\theta) = \begin{bmatrix} \frac{\partial^2 \ell(\theta; x)}{\partial \alpha^2} & \frac{\partial^2 \ell(\theta; x)}{\partial \tau \partial \alpha} & \frac{\partial^2 \ell(\theta; x)}{\partial \lambda \partial \alpha} \\ \frac{\partial^2 \ell(\theta; x)}{\partial \alpha \partial \tau} & \frac{\partial^2 \ell(\theta; x)}{\partial \tau^2} & \frac{\partial^2 \ell(\theta; x)}{\partial \lambda \partial \tau} \\ \frac{\partial^2 \ell(\theta; x)}{\partial \alpha \partial \lambda} & \frac{\partial^2 \ell(\theta; x)}{\partial \tau \partial \lambda} & \frac{\partial^2 \ell(\theta; x)}{\partial \lambda^2} \end{bmatrix}$$

For large sample sizes, the MLE parameter vector is approximately normal with mean θ and variance-covariance matrix which is the inverse of J . This is helpful when constructing confidence intervals for the parameters estimated by the MLEs.

3.1.6 Ordering

We provide here the stochastic ordering of two independent random variables X and Y . Let $X \sim \text{GiPL}(\alpha_1, \tau_1, \lambda_1)$ and $Y \sim \text{GiPL}(\alpha_2, \tau_2, \lambda_2)$. Then, define the following types of the ordering of X and Y :

Stochastic Ordering: X is at most Y in stochastic order, denoted by $X \leq_S Y$, if $F_X(x) \geq F_Y(y)$, for all x , where F_X and F_Y are the cumulative distribution functions of X and Y , respectively.

Likelihood-Ratio Ordering: X is at most Y in likelihood ratio order, denoted by $X \leq_L Y$, if $f_X(x) / f_Y(x)$ is decreasing, for all x , where f_X and f_Y are the probability density function of X and Y , respectively.

Hazard Ordering: X is at most Y in hazard order, denoted by $X \leq_h Y$, if $h_X(x) \geq h_Y(x)$, for all x , where h_X and h_Y are the hazard function of X and Y , respectively.

The following are the known series of implications for the ordering:

$$X \leq_L Y \rightarrow X \leq_h Y \rightarrow X \leq_S Y.$$

Now, using Equation 5, we get

$$\begin{aligned} \ln \frac{f_X(x)}{f_Y(y)} &= \ln \frac{\Gamma(\alpha_2)}{\Gamma(\alpha_1)} + (\alpha_1 - 1) \ln \left\{ -\ln \left[1 - \left(\frac{(\lambda_1 x)^{\tau_1}}{1 + (\lambda_1 x)^{\tau_1}} \right)^{\tau_1} \right] \right\} \\ &\quad - (\alpha_2 - 1) \ln \left\{ -\ln \left[1 - \left(\frac{(\lambda_2 x)^{\tau_2}}{1 + (\lambda_2 x)^{\tau_2}} \right)^{\tau_2} \right] \right\} + 2(\ln \tau_1 - \ln \tau_2) \\ &\quad + \tau_1^2 (\ln \lambda_1 + \ln x) - \tau_2^2 (\ln \lambda_2 + \ln x) - (\tau_1 + 1) \ln [1 + (\lambda_1 x)^{\tau_1}] \\ &\quad + (\tau_2 + 1) \ln [1 + (\lambda_2 x)^{\tau_2}] \end{aligned}$$

Consequently,

$$\frac{d}{dx} \left[\ln \frac{f_X(x)}{f_Y(y)} \right] = (\alpha_1 - 1) \left\{ -\ln \left[1 - \left(\frac{(\lambda_1 x)^{\tau_1}}{1 + (\lambda_1 x)^{\tau_1}} \right)^{\tau_1} \right] \right\}^{-1} \frac{\tau_1^2 (\lambda_1 x)^{\tau_1}}{x [1 + (\lambda_1 x)^{\tau_1}]^{\tau_1 + 1}}$$

$$\begin{aligned}
 & -(\alpha_2 - I) \left\{ -\ln \left[I - \left(\frac{(\lambda_2 x)^{\tau_2}}{I + (\lambda_2 x)^{\tau_2}} \right)^{\tau_2} \right] \right\}^{-I} \frac{\tau_2^2 (\lambda_2 x)^{\tau_2}}{x [I + (\lambda_2 x)^{\tau_2}]^{\tau_2 + I}} \\
 & + \frac{\tau_1^2 - \tau_2^2}{x} - \frac{\tau_1 (\tau_1 + I)}{x} \frac{(\lambda_1 x)^{\tau_1}}{I + (\lambda_1 x)^{\tau_1}} + \frac{\tau_2 (\tau_2 + I)}{x} \frac{(\lambda_2 x)^{\tau_2}}{I + (\lambda_2 x)^{\tau_2}}
 \end{aligned}$$

If $\alpha_1 > \alpha_2$, $\tau_1 = \tau_2$, and $\lambda_1 < \lambda_2$, then observe that

1. $\frac{\tau_1^2 - \tau_2^2}{x} - \frac{\tau_1 (\tau_1 + I)}{x} \frac{(\lambda_1 x)^{\tau_1}}{I + (\lambda_1 x)^{\tau_1}} + \frac{\tau_2 (\tau_2 + I)}{x} \frac{(\lambda_2 x)^{\tau_2}}{I + (\lambda_2 x)^{\tau_2}} > 0$
2. $\frac{(\alpha_1 - I) \left\{ -\ln \left[I - \left(\frac{(\lambda_1 x)^{\tau_1}}{I + (\lambda_1 x)^{\tau_1}} \right)^{\tau_1} \right] \right\}^{-I}}{(\alpha_2 - I) \left\{ -\ln \left[I - \left(\frac{(\lambda_2 x)^{\tau_2}}{I + (\lambda_2 x)^{\tau_2}} \right)^{\tau_2} \right] \right\}^{-I}} > I$

Furthermore, if $0 < t_1 \leq \frac{I + \sqrt{5}}{2}$, then

$$\begin{aligned}
 \frac{\tau_1^2 (\lambda_1 x)^{\tau_1}}{x [I + (\lambda_1 x)^{\tau_1}]^{\tau_1 + I}} &= \tau_1^2 (\lambda_1 x)^{\tau_1^2 - \tau_1 - I} \left(\frac{(\lambda_1 x)^{\tau_1}}{I + (\lambda_1 x)^{\tau_1}} \right)^{\tau_1 + I} \\
 &> \tau_1^2 (\lambda_2 x)^{\tau_1^2 - \tau_1 - I} \left(\frac{(\lambda_2 x)^{\tau_1}}{I + (\lambda_2 x)^{\tau_1}} \right)^{\tau_1 + I} = \frac{\tau_1^2 (\lambda_2 x)^{\tau_1}}{x [I + (\lambda_2 x)^{\tau_1}]^{\tau_1 + I}}
 \end{aligned}$$

From the three inequalities above, we conclude that $\frac{d}{dx} \left[\ln \frac{f_X(x)}{f_Y(x)} \right] > 0$. Therefore, $f_X(x)/f_Y(x)$ is increasing, for all $x > 0$, which implies that $f_Y(x)/f_X(x)$ is decreasing. This result and its consequences are summarized in the following theorem.

Theorem 4. Let $X \sim \text{GiPL}(\alpha_1, \tau_1, \lambda_1)$ and $Y \sim \text{GiPL}(\alpha_2, \tau_2, \lambda_2)$. If $\alpha_1 > \alpha_2$, $\tau_1 = \tau_2$, $\lambda_1 < \lambda_2$, and $0 < t_1 \leq \frac{I + \sqrt{5}}{2}$, then $Y \leq_L X$. Consequently, $Y \leq_h X$ which implies that $Y \leq_S X$.

3.2 Parameter Estimation

This section presents a simulation study to test the performance of the MLEs. It is followed by an application of the model given a data set. Furthermore, it is compared with various recently developed models.

3.2.1 Simulation Study

Using the quantile function given in Equation 7, sample data of sizes $n = 25, 50, 75, 100, 150$ and 200 were generated. The generation of the sample was

repeated 10,000 times. Summarized in Table 2 are the mean estimates, biases and RMSE of the parameters with parameter values $(\alpha, \tau, \lambda) = (1.4, 2.5, 3.7)$ and $(\alpha, \tau, \lambda) = (2.5, 3.5, 5.5)$ corresponding to the sample sizes.

It can be observed that the estimates approach the true value of the parameters as the number of sample sizes increased. Moreover, the biases were getting close to zero. It was also evident that the RMSEs were decreasing whenever the sample size was increasing.

Table 2. Mean estimates, biases and RMSEs of simulations with parameter values for sample sizes 25, 50, 75, 100, 150 and 200

Sample size (n)	$(\alpha, \tau, \lambda) = (1.4, 2.5, 3.7)$				Sample size (n)	$(\alpha, \tau, \lambda) = (2.5, 3.5, 5.5)$			
	Parameters	Mean	Bias	RMSE		Parameters	Mean	Bias	RMSE
25	α	1.4179	-0.0179	0.1989	25	α	2.4891	-0.0284	0.2078
	τ	2.5882	-0.0882	0.3765		τ	3.3214	-0.1125	0.275
	λ	3.7035	-0.0035	0.3966		λ	5.6264	-0.0059	0.4367
50	α	1.4115	-0.0115	0.1425	50	α	2.4911	-0.0225	0.144
	τ	2.541	-0.041	0.2575		τ	3.3205	-0.0712	0.2093
	λ	3.697	0.003	0.2851		λ	5.6229	-0.0045	0.3522
75	α	1.4081	-0.0081	0.1156	75	α	2.4923	-0.0192	0.1409
	τ	2.5244	-0.0244	0.2052		τ	3.4387	-0.0553	0.1545
	λ	3.6976	0.0024	0.229		λ	5.6032	-0.0041	0.2657
100	α	1.4079	-0.0079	0.1017	100	α	2.4944	-0.0179	0.107
	τ	2.5164	-0.0164	0.1779		τ	3.4553	-0.0411	0.1354
	λ	3.6953	0.0046	0.2		λ	5.5871	-0.0035	0.2487
150	α	1.4045	-0.0046	0.0825	150	α	2.4969	-0.0166	0.0952
	τ	2.5092	-0.0092	0.1411		τ	3.482	-0.0358	0.116
	λ	3.6984	0.0016	0.1628		λ	5.5424	-0.0019	0.243
200	α	1.4029	-0.0029	0.0725	200	α	2.4993	-0.0062	0.0733
	τ	2.5078	-0.0078	0.1198		τ	3.4827	-0.0167	0.0719
	λ	3.6999	0.000095	0.143		λ	5.5312	-0.0004	0.1919

3.2.2 Application to Real-Time Data

This subsection shows the superiority of the GiPL(α, τ, λ) distribution among other recently developed distribution models: generalized Log Moyal (GlogM μ, σ) (Bhati and Ravi, 2018), Marshall-Olkin Kappa distribution (MOK($\alpha, \beta, \theta, \delta$)) (Javed *et al.*, 2019), gamma generalized normal distribution (GNG(μ, σ, s, a)) (Cordeiro *et al.*, 2019), Marshall-Olkin length biased exponential distribution MOLBE(γ, β) (ul Haq *et al.*, 2019), Marshall-Olkin extended inverted Kumaraswamy distribution (MOEIK λ, α, β) (Usman and ul Haq, 2020), Marshall-Olkin power Lomax distribution MOPL($\alpha, \beta, \gamma, \lambda$) (ul Haq *et al.*, 2020) and Marshall-Olkin inverted Nadarajah-Haghighi distribution MOINH(α, β, γ) (Raffiq *et al.*, 2020). These are heavy-tailed distributions that can be used to model data sets that deviate from the normal distribution in terms of numerous statistical properties.

As observed in the probability-probability plot in Figure 6, there was a significant variation over the right tail. This was expected as the statistics presented in the previous section to describe the distribution of the data suggested that the distribution was positively skewed and heavy-tailed indicating the need to use a heavy-tailed distribution to model the data set.

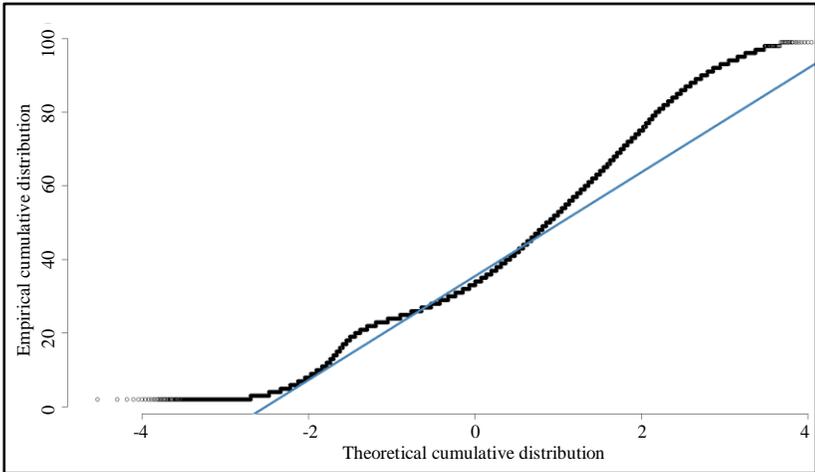


Figure 6. Probability-probability plot given the COVID-19 data set

The measures for model selection used to assess the superiority of the $GiPL(\alpha, \tau, \lambda)$ were log-likelihood (log-LL), Akaike Information Criterion (AIC), corrected Akaike Information Criterion (AIC_c), and Bayesian Information Criterion (BIC). The goal was to compare models and determine a model with the highest log-LL, and lowest AIC, AIC_c , and BIC. The maxLik package in R was employed with Nelder-Mead optimization to estimate the MLEs of the parameters, thanks to Henningsen and Toomet (2010).

Given the data set on the number of COVID-19 infected individuals per age, summarized in Table 3 are the estimates of the parameters of $GiPL(\alpha, \tau, \lambda)$, $GlogM(\mu, \sigma)$, $MOK\alpha, \beta, \theta, \delta$, $GGN(\mu, \sigma, s, a)$, $MOEIK(\lambda, \alpha, \beta)$, $MOLBE(\gamma, \beta)$, $MOPL(\alpha, \beta, \gamma, \lambda)$ and $MOINH(\alpha, \beta, \gamma)$.

Furthermore, the corresponding values of log-LL, AIC, AIC_c , and BIC are also given. The $GiPL(\alpha, \tau, \lambda)$ dominated all the other models considered; that is, $GiPL(\alpha, \tau, \lambda)$ had the highest log-LL and the lowest AIC, AIC_c and BIC. Therefore, $GiPL(\alpha, \tau, \lambda)$ gives the best fit compared with the other models.

Table 3. Parameter estimates and measures for model selection using the data set on the number of individuals per age who were infected with COVID-19 in NCR, Philippines from April 2020 to March 2022

Model	Parameter estimation (Standard error)	Measures for model selection			
		log-ll	AIC	BIC	AIC _c
GiPL(α, τ, λ)	$\hat{\alpha} = 0.0163$ (0.0004)	-752122.4	1504251	1504267	1504251
	$\hat{\tau} = 3.376$				
	(5.4×10^{-6})				
	$\hat{\lambda} = 0.0020$ (1.6×10^{-8})				
GlogM(μ, σ)	$\hat{\mu} = 0.7494$ (0.0012)	-870337	1740678	1740689	1740678
	$\hat{\sigma} = 16.5010$ (0.0356)				
MOK($\alpha, \beta, \theta, \delta$)	$\hat{\alpha} = 2.0098$ (0.0080)	-760724.9	1521458	1521479	1521458
	$\hat{\beta} = 22.9569$ (0.0382)				
	$\hat{\theta} = 2.4713$ (0.0131)				
	$\hat{\delta} = 15.8070$ (0.1268)				
GGN(μ, σ, s, a)	$\hat{\mu} = 10.14$ (1.9×10^{-5})	-756556.3	1513121	1513142	1513121
	$\hat{\sigma} = 0.3444$ (9.9×10^{-6})				
	$\hat{s} = 0.5880$ (9.3×10^{-6})				
	$\hat{a} = 16.61$ (0.0088)				
MOEIK(λ, α, β)	$\hat{\lambda} = 1.1834$ (0.0047)	-1189690	2379386	2379402	2379386
	$\hat{\alpha} = 1.1098$ (0.0024)				
	$\hat{\beta} = 26.4833$ (0.1518)				
MOLBE(γ, β)	$\gamma = 2.97 \times 10^{-3}$ (1.1×10^{-6})	-788627.4	1577259	1577269	1577259
	$\beta = 0.0015$ (0.0150)				
MOPL($\alpha, \beta, \gamma, \lambda$)	$\alpha = 0.0064$ (9.975×10^{-6})	-863759.7	1727527	1727549	1727527
	$\beta = 154.3$ (0.0089)				
	$\gamma = 19.91$ (0.1169)				
	$\lambda = 3.639$ (0.0003)				
MOINH(α, β, γ)	$\alpha = 3.8546$ (0.0313)	-847350.5	1694707	1694723	1694707
	$\beta = 1.2028$ (0.0119) $\gamma = 7.2761$ (0.0316)				

In Figure 7 the COVID-19 data set histogram along with the fitted probability density function using the estimates for $GiPL(\alpha, \tau, \lambda)$ in Table 4 show that there were some small overestimation and underestimation over the left and right tails, respectively. But despite this, a good picture of the fitted model compared to the histogram of the data set is observed.

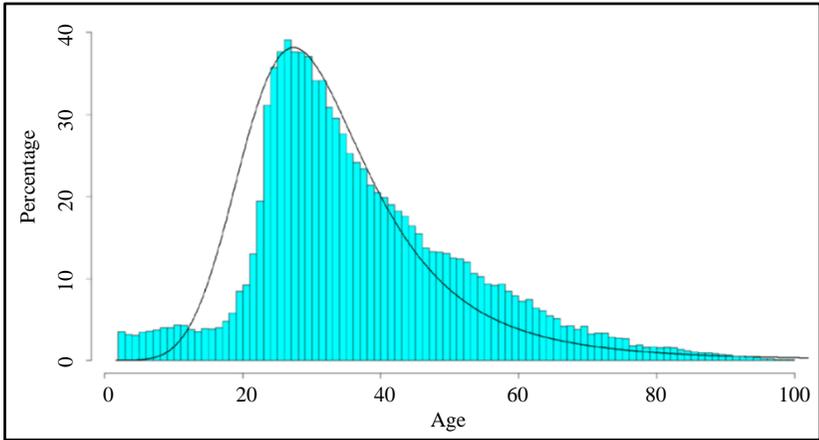


Figure 7. Histogram with fitted probability density function

Figure 8 shows the empirical cumulative distribution function with fitted $GiPL$ cumulative distribution function using the estimates for $GiPL(\alpha, \tau, \lambda)$ in Table 3. It provides an illustration of how adequate the $GiPL(\alpha, \tau, \lambda)$ is used to model the data set in terms of cumulative distribution.

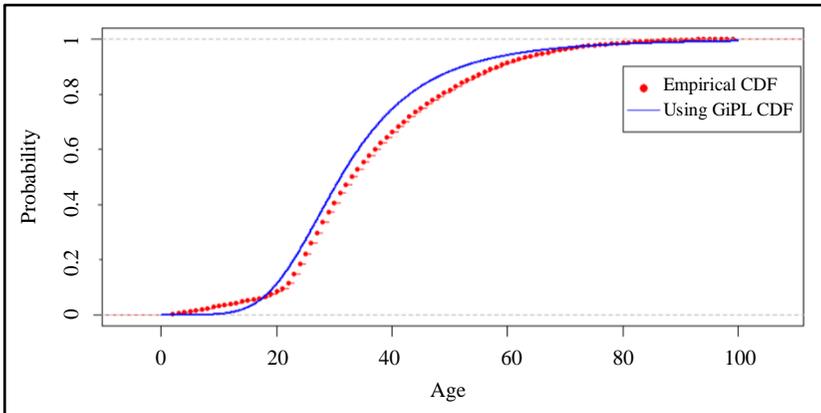


Figure 8. Empirical distribution with fitted $GiPL$ distribution

In summary, Figure 9 presents the log-LL plot to illustrate the behavior of the log-LL over the neighborhood of the two-parameter estimates having other parameters constant.

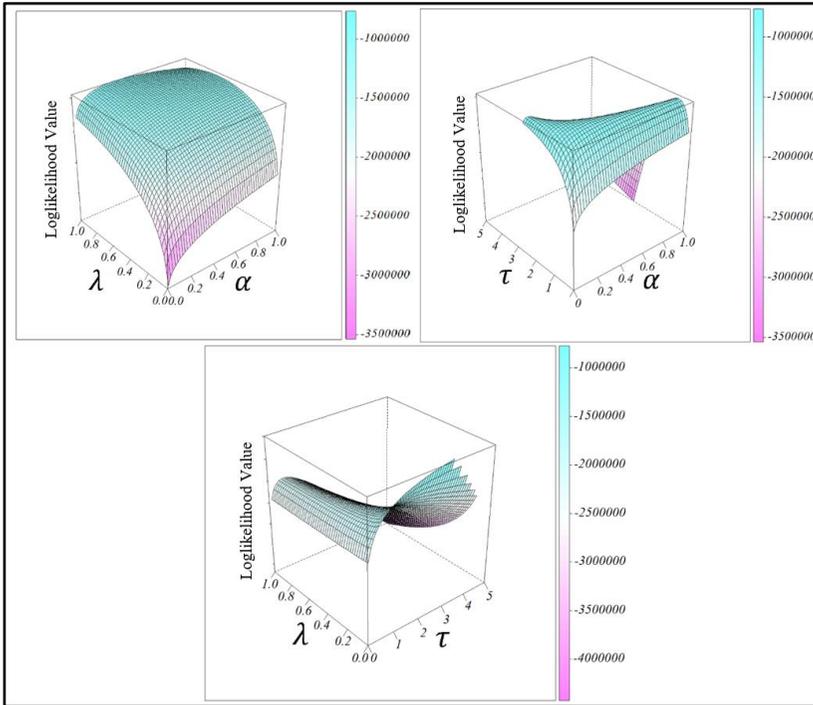


Figure 9. Loglikelihood plot for $\text{GiPL}(\alpha, \tau, \lambda)$

4. Conclusion and Recommendation

A three-parameter distribution model called the GiPL distribution, denoted by $\text{GiPL}(\alpha, \theta, \tau)$, was proposed. The development made use of the inverse paralogistic model with a gamma family generator. Introduced together were the probability density and cumulative distribution functions and quantile functions of $\text{GiPL}(\alpha, \theta, \tau)$. Properties such as some measures for reliability analysis, the raw moment and moment-generating function, partial moments, order statistics, log-likelihood functions, the Renyi entropy, the stochastic, likelihood-ratio and the hazard ordering were presented. The performance of the parameters was studied using the MLEs. The mean estimates, biases, and RMSEs showed that as the sample size increases, the mean estimates

approached the true value of the parameters initially set, the biases were close to zero, and the RMSEs relatively decreased. Finally, as the data description implied the use of a heavy-tailed distribution model because the data set was positively skewed and the kurtosis relatively deviated from the normal, it was used to show the superiority of the GiPL(α, θ, τ) over the recently developed models. Particularly, measures such as log-LL value, AIC, BIC and AIC_c for model selection suggested that GiPL(α, θ, τ) dominated the seven models used because of having the highest log-LL and the lowest value of AIC, BIC and AIC_c. Among all these models, the MOEIK(λ, α, β) was the least exceptional.

5. References

- Bhati, D., & Ravi, S. (2018). On generalized log-Moyal distribution: A new heavy tailed size distribution. *Insurance: Mathematics and Economics*, 79, 247-259. <https://doi.org/10.1016/j.insmatheco.2018.02.002>
- Biostatistics Research Cluster. (2022). Philippine COVID-19 case information dashboard. Retrieved from <https://datastudio.google.com/s/rWhJ8gT9g7U>
- Cordeiro, G.M., Cintra, R.J., Rego, L.C., & Nascimento, A.D. (2019). The gamma generalized normal distribution: A descriptor of SAR imagery. *Journal of Computational and Applied Mathematics*, 347, 257-72. <https://doi.org/10.1016/j.cam.2018.07.045>
- De Andrade, T.A., Fernandez, L.M., Gomes-Silva, F., & Cordeiro, G.M. (2017). The gamma generalized pareto distribution with applications in survival analysis. *International Journal of Statistics and Probability*, 6, 141. <https://doi.org/10.5539/ijsp.v6n3p141>
- Guerra, R., Ramirez, F., & Cordeiro, G. (2017). The gamma Burr XII distributions: Theory and applications. *Journal of Data Science*, 15, 467-494. [https://doi.org/10.6339/JDS.201707_15\(3\).0006](https://doi.org/10.6339/JDS.201707_15(3).0006)
- Henningsen, A., & Toomet, O. (2010). maxLik: A package for maximum likelihood estimation in R. *Computational Statistics*, 26, 443-458. <https://doi.org/10.1007/s00180-010-0217-1>
- Javed, M., Nawaz, T., & Irfan, M. (2019). The Marshall-Olkin Kappa distribution: Properties and applications. *Journal of King Saud University - Science*, 31, 684-691. <https://doi.org/10.1016/j.jksus.2018.01.001>
- Kantar, Y.M., Usta, I., Arik, I., & Yenilmez, I. (2018). Wind speed analysis using the Extended Generalized Lindley Distribution. *Renewable Energy*, 118, 1024-1030. <https://doi.org/10.1016/j.renene.2017.09.053>

Nawaz, T., Hussain, S., Ahmad, T., Naz, F., & Abid, M. (2020). Kumaraswamy generalized Kappa distribution with application to stream flow data. *Journal of King Saud University - Science*, 32, 172-182. <https://doi.org/10.1016/j.jksus.2018.04.005>

Raffiq, G., Dar, I. S., ul Haq, M.A., & Ramos, E. (2020). The Marshall–Olkin inverted Nadarajah–Haghighi distribution: Estimation and applications. *Annals of Data Science*, 9, 1323-1338. <https://doi.org/10.1007/s40745-020-00297-7>

R Core Team. (2023). A language and environment for statistical computing [Computer software]. Vienna, Austria: R Foundation for Statistical Computing.

ul Haq, M.A., Rao, G.S., Albassam, M., & Aslam, M. (2020). Marshall–Olkin Power Lomax distribution for modeling of wind speed data. *Energy Reports*, 6, 1118-1123. <https://doi.org/10.1016/j.egyr.2020.04.033>

ul Haq, M.A., Usman, R.M., Hashimi, S., & Al-Omeri, A.I. (2019). The Marshall–Olkin length-biased exponential distribution and its applications. *Journal of King Saud University - Science*, 31(2), 246-251. <https://doi.org/10.1016/j.jksus.2017.09.006>

Usman, R.M., & Ilyas, M. (2020). The power Burr type X distribution: Properties, regression modeling and applications. *Punjab University Journal of Mathematics*, 52(8), 27-44.

Usman, R.M., Handique, L., & Chakraborty, S. (2019). Some aspects of the odd log-logistic Burr X distribution with applications in reliability data modeling. *International Journal of Applied Mathematics and Statistics*, 58(1).

Usman, R., & ul Haq, M. (2020). The Marshall–Olkin extended inverted Kumaraswamy distribution: Theory and applications. *Journal of King Saud University - Science*, 356-365. <https://doi.org/10.1016/j.jksus.2018.05.021>

Zografos, K., & Balakrishnan, N. (2009). On families of beta- and generalized gamma-generated distributions and associated inference. *Statistical Methodology*, 6, 344-362. <https://doi.org/10.1016/j.stamet.2008.12.003>