Sains Malaysiana 54(3)(2025): 927-941
http://doi.org/10.17576/jsm-2025-5403-23

A New Gompertz-Three-Parameter-Lindley Distribution for Modeling Survival Time Data

(Taburan Gompertz-Tiga-Parameter-Lindley Baharu untuk Memodelkan Data Masa Kemandirian)

FEI LIANG[1], HEZHI LU[1,2,*] & YUHANG XI[1]

[1]*School of Economics and Statistics, Guangzhou University, Guangzhou 510006, China*
[2]*Lingnan Research Academy of Statistical Science, Guangzhou University, Guangzhou 510006, China*

Received: 20 May 2024/Accepted: 2 December 2024

ABSTRACT

In this paper, a new survival distribution is introduced. It is a mixture of the Gompertz distribution and three-parameter-Lindley distribution. The statistical properties of the proposed distribution including the shape properties, cumulative distribution, quantile functions, moment generating function, failure rate function, mean residual function, and stochastic orders are studied. Moreover, a new regression model based on the proposed distribution is presented. Maximum likelihood estimators (MLEs) of unknown parameters are obtained via differential evolution algorithms, and simulation studies are conducted to evaluate the consistency of the MLEs. Finally, the proposed model and its regression model are applied to a real dataset and compared with other well-known models, demonstrating their superior performance, particularly for heavy-tailed data.

Keywords: Differential evolution algorithm; Gompertz-Lindley distribution; maximum likelihood estimation; regression model; structural property

ABSTRAK

Dalam kertas ini, suatu taburan survival baharu diperkenalkan. Ia adalah campuran taburan Gompertz dan taburan tiga parameter-Lindley. Sifat statistik bagi taburan yang dicadangkan termasuk sifat bentuk, taburan kumulatif, fungsi kuantil, fungsi penjanaan momen, fungsi kadar kegagalan, fungsi baki min dan susunan stokastik dikaji. Selain itu, model regresi baharu berdasarkan pengedaran yang dicadangkan dibentangkan. Anggaran kebolehjadian maksimum (MLE) bagi parameter yang tidak diketahui diperoleh melalui algoritma evolusi pembezaan, dan kajian simulasi dijalankan untuk menilai ketekalan MLE. Akhir sekali, model yang dicadangkan dan model regresinya digunakan pada set data sebenar dan dibandingkan dengan model terkenal lain, menunjukkan prestasi unggul mereka, terutamanya untuk data berat.

Kata kunci: Algoritma evolusi berbeza; anggaran kebolehjadian maksimum; model regresi; sifat struktur; taburan Gompertz-Lindley

INTRODUCTION

The Gompertz distribution (Gavrilov & Gavrilova 2001) is commonly used in diverse areas such as actuarial science, survival analysis, and reliability engineering, allowing researchers and practitioners to understand and model complex phenomena related to mortality rates, time-to-event data, and system failures. In general, let X be a random variable following the Gompertz distribution with frailty parameter θ and scale parameter λ, with a conditional probability density function (pdf):

$$f(x \mid \theta,\lambda) = \lambda\theta \exp\{\lambda x - \theta(e^{\lambda x} - 1)\}, x > 0, \theta, \lambda > 0. \quad (1)$$

The maximum likelihood methods of the parameters θ and λ based on random samples and progressively type-II censored samples were studied by Lenart (2014) and Ghitany, Alqallaf and Balakrishnan (2014), respectively. Moreover, Lenart and Missov (2016) studied the goodness-of-fit tests for the Gompertz distribution.

In real data analysis, the collected data are often heterogeneous, with various parts of the data following different distributions. In this case, the traditional single model may not meet practical needs. To obtain a more flexible distribution, well-established distribution expansion techniques such as the transmutation map method, distribution weighting method and mixture model method have been discussed. To date, numerous mixed Gompertz distribution models, such as the beta-Gompertz (Jafari, Tahmasebi & Alizadeh 2014), beta generalized Gompertz (Benkhelifa 2017), odd generalized exponential Gompertz (El-Damcese et al. 2015), generalized Gompertz-generalized Gompertz (Boshi, Abid & Al-Noor 2019), Gompertz-Lindley (Ghitany et al. 2019), odd log-logistic generalized Gompertz (Alizadeh et al. 2020), Marshall-Olkin Gompertz (Eghwerido, Ogbo & Omotoye 2021), generalized Gompertz (Jayakumar & Shabeer 2022), Gamma-Gompertz (Shama et al. 2022), and Gompertz-two-parameter-Lindley distributions (Ou, Lu & Kong 2022) have been introduced.

Since the Lindley distribution (LD) proposed by Lindley (1958) is a useful tool in statistical modeling for data with nonnegative values and decreasing hazard rates, in recent years, several new mixture models have been obtained by mixing the Gompertz distribution and Lindley-type distributions, which have better adaptability in capturing the heterogeneity of the data. Suppose that the frailty parameter θ in (1) follows an LD (Ghitany, Atieh & Nadarajah 2008) with shape parameter α, with pdf.

$$h(\theta)=\frac{\alpha^2}{\alpha+1}(1+\theta)\mathrm{e}^{-\alpha\theta},\ \theta>0,\alpha>0. \tag{2}$$

Then, the unconditional pdf of the Gompertz–Lindley (GL) distribution proposed by Ghitany et al. (2019) is given by

$$f(x)=\frac{\lambda\alpha^2e^{\lambda x}(e^{\lambda x}+\alpha+1)}{(\alpha+1)(e^{\lambda x}+\alpha-1)^3},\ x>0,\alpha,\lambda>0. \tag{3}$$

Moreover, Shanker and Sharma (2013) proposed a generalized Lindley distribution with shape parameters α and β as follows:

$$k(\theta)=\frac{\alpha^2}{\alpha+\beta}(1+\beta\theta)\mathrm{e}^{-\alpha\theta},\theta>0,\alpha>0,\beta>-\alpha. \tag{4}$$

By using the two-parameter LD (4) to replace the LD in (2), Ou, Lu and Kong (2022) derived a generalized three-parameter GL distribution as follows:

$$f(x)=\frac{\lambda\alpha^2e^{\lambda x}(e^{\lambda x}+\alpha+2\beta-1)}{(\alpha+\beta)(e^{\lambda x}+\alpha-1)^3},x>0,\alpha,\lambda>0,\beta\geq 0. \tag{5}$$

To date, a new generalized three-parameter LD method proposed by Shanker et al. (2017) and its extensions (Shanker, Shukla & Mishra 2017) and applications (Al-Omari, Ciavolino & Al-Nasser 2020; Thamer & Zine 2023) have attracted increased research interest. The pdf of the model with parameters α,β and η are defined as follows:

$$g(\theta)=\frac{\alpha^2}{\alpha\beta+\eta}(\beta+\eta\theta)e^{-\alpha\theta}, \quad \theta>0,\alpha,\eta>0,\alpha\beta+\eta>0. \tag{6}$$

Note that the pdf of the three-parameter LD in Shanker et al. (2017) and Xi, Lu and Liang (2024) may not satisfy the regularization condition of probability. To solve this issue, we use the parameter space $\Omega=\{\beta:\beta\geq 0\}$ of β to replace the parameter space $\Omega'=\{\beta:\beta>-\eta/\alpha\}$ of β in (6).

To obtain a new more flexible and adaptable model, we use the new modified three-parameter LD in (6) to replace the LD in (2) when combining the Gompertz distribution with the LD. The corresponding unconditional pdf of X is given by

$$\begin{aligned}f(x)&=\frac{\lambda\alpha^2e^{\lambda x}}{\alpha\beta+\eta}\int_0^\infty(\beta\theta+\eta\theta^2)\exp\{-\theta(e^{\lambda x}+\alpha-1)\}d\theta\\&=\frac{\lambda\alpha^2e^{\lambda x}(\beta(e^{\lambda x}+\alpha-1)+2\eta)}{(\alpha\beta+\eta)(e^{\lambda x}+\alpha-1)^3},x>0,\alpha,\eta,\lambda>0,\beta\geq 0.\end{aligned} \tag{7}$$

We refer to the new random variable X as the Gompertz-Three-Parameter-Lindley distribution with shape parameters α, β and η and scale parameter λ, which are denoted by GTHPL$(\alpha,\beta,\eta,\lambda)$. Figure 1 displays the pdfs of the GTHPL distributions for various values of α,β,η and λ. For example, with fixed $\beta=2,\eta=1$ and $\lambda=2$, increasing α leads to a distribution that accommodates a larger right tail. Hence, the proposed model is well suited for datasets where there is a significant right tail. Moreover, the structural properties and associated inference of the proposed model are considered.

STRUCTURAL PROPERTIES OF THE MODEL

SHAPE PROPERTIES

Theorem 2.1 For all $\alpha,\eta,\lambda>0,\beta\geq 0$, the PDF $f(x)$ of GTHPL$(\alpha,\beta,\eta,\lambda)$ has the following shape properties:

(i) if $0<\alpha\leq 2$, $f(x)$ is decreasing;
(ii) unimodal if $\alpha\geq 3$;
(iii) if $\eta\geq\frac{\alpha\beta(2-\alpha)}{2(\alpha-3)}$ $\left(\eta<\frac{\alpha\beta(2-\alpha)}{2(\alpha-3)}\right)$ and $2<\alpha<3$, then $f(x)$ is decreasing (unimodal).

Proof The first derivative of $f(x)$ is given by

$$f'(x)=\frac{\lambda^2\alpha^2e^{\lambda x}}{(\alpha\beta+\eta)(e^{\lambda x}+\alpha-1)^4}\left(-\beta\left(e^{\lambda x}\right)^2-4\eta e^{\lambda x}+(\alpha-1)(\alpha\beta-\beta+2\eta)\right).$$

Let $\xi(t)=-\beta t^2-4\eta t+(\alpha-1)(\alpha\beta-\beta+2\eta)$, where $t=e^{\lambda x}$. According to Descartes' rule of signs, since $\xi(t)$ is a unimodal function in t, the expression $\xi(1)=\beta\alpha(\alpha-2)+2\eta(\alpha-3)$ is negative (changing sign from positive to negative) when $0<\alpha\leq 2$ $(\alpha\geq 3)$. Furthermore, consider the function $\varsigma(\eta)=\beta\alpha(\alpha-2)+2\eta(\alpha-3)$, where $2<\alpha<3$.

$\eta = \frac{\alpha\beta(2-\alpha)}{2(\alpha-3)}$ is the solution of equation $\varsigma(\eta) = 0$. Since $\varsigma(\eta)$ is a decreasing function in η, $\xi(t)$ is negative (sign change from positive to negative) when $\eta \geq \frac{\alpha\beta(2-\alpha)}{2(\alpha-3)}\left(\eta < \frac{\alpha\beta(2-\alpha)}{2(\alpha-3)}\right)$ and $2 < \alpha < 3$. The proof is complete.

Remark 2.1 Let $\beta = 1$, GTHPL$(\alpha, \beta, \eta, \lambda)$ reduces to the Gompertz-two-parameter-Lindley distribution (Ou, Lu & Kong 2022). Moreover, for $\beta = 1$, $\eta = 0$ and $\eta = 1$, the GTHPL distribution reduces to the extended exponential distribution (Marshall & Olkin 1997) and Gompertz–Lindley distribution (Ghitany et al. 2019), respectively.

Remark 2.2 For $\lambda > 0$ and $\beta \geq 0$, if $0 < \alpha \leq 2$ or $2 < \alpha < 3$ and $\eta < \frac{\alpha\beta(2-\alpha)}{2(\alpha-3)}$, the mode of x is equal to 0. On the other hand, if $\alpha \geq 3$ or $2 < \alpha < 3$ and $\eta < \frac{\alpha\beta(2-\alpha)}{2(\alpha-3)}$, then the mode of x is the solution of $f'(x) = 0$. Given that $t_0 = e^{\lambda x_0} = \frac{-2\eta + \left(4\eta^2 + \beta(\alpha-1)(\beta(\alpha-1)+2\eta)\right)^{\frac{1}{2}}}{\beta}$ is the solution of the quadratic equation $-\beta t_0^2 - 4\eta t_0 + (\alpha-1)(\alpha\beta - \beta + 2\eta) = 0$, then the mode of x is

$$\frac{\ln\left(-2\eta + \left(4\eta^2 + \beta(\alpha-1)(\beta(\alpha-1)+2\eta)\right)^{\frac{1}{2}}\right) - \ln\beta}{\lambda}.$$

CUMULATIVE DISTRIBUTION AND QUANTILE FUNCTIONS

The cumulative distribution function (cdf) of GTHPL$(\alpha, \beta, \eta, \lambda)$ is given as follows:

$$\begin{aligned} F(x) = P(x \leq x) &= \int_0^x \frac{\lambda\alpha^2 e^{\lambda u}\left(\beta\left(e^{\lambda u}+\alpha-1\right)+2\eta\right)}{(\alpha\beta+\eta)\left(e^{\lambda u}+\alpha-1\right)^3} du \\ &= \frac{\alpha^2\beta}{\alpha\beta+\eta}\int_0^x \frac{1}{\left(e^{\lambda u}+\alpha-1\right)^2} d\left(e^{\lambda u}+\alpha-1\right) \\ &\quad + \frac{2\eta\alpha^2}{\alpha\beta+\eta}\int_0^x \frac{1}{\left(e^{\lambda u}+\alpha-1\right)^3} d\left(e^{\lambda u}+\alpha-1\right) \\ &= 1 - \frac{\alpha^2\left(\beta\left(e^{\lambda x}+\alpha-1\right)+\eta\right)}{(\alpha\beta+\eta)\left(e^{\lambda x}+\alpha-1\right)^2}, x > 0, \alpha, \eta, \lambda > 0, \beta \geq 0. \end{aligned} \quad (8)$$

Figure 2 displays the cdfs of the GTHPL distributions for various values of α, β, η and λ.

The qth quantile x_q of GTHPL$(\alpha, \beta, \eta, \lambda)$ is determined as the solution of $q = F(x_q)$, where $0 < q < 1$. Consequently, $x_q = F^{-1}(q)$ is equivalent to

$$x_q = \lambda^{-1}\ln\left(\frac{\alpha\left(\left(\alpha^2\beta^2 + 4(1-q)(\alpha\beta+\eta)\eta\right)^{\frac{1}{2}} + \alpha\beta\right)}{2(\alpha\beta+\eta)(1-q)} - \alpha + 1\right). \quad (9)$$

Hence, the quantiles of GTHPL$(\alpha, \beta, \eta, \lambda)$ are given as follows:

$$Q_1 = \lambda^{-1}\ln\left(\frac{2\alpha\left(\left(\alpha^2\beta^2 + 3(\alpha\beta+\eta)\eta\right)^{\frac{1}{2}} + \alpha\beta\right)}{3(\alpha\beta+\eta)} - \alpha + 1\right), \quad (10)$$

$$Q_2 = \lambda^{-1}\ln\left(\frac{2\alpha\left(\left(\alpha^2\beta^2 + 2(\alpha\beta+\eta)\eta\right)^{\frac{1}{2}} + \alpha\beta\right)}{2(\alpha\beta+\eta)} - \alpha + 1\right), \quad (11)$$

$$Q_3 = \lambda^{-1}\ln\left(\frac{2\alpha\left(\left(\alpha^2\beta^2 + (\alpha\beta+\eta)\eta\right)^{\frac{1}{2}} + \alpha\beta\right)}{(\alpha\beta+\eta)} - \alpha + 1\right). \quad (12)$$

FAILURE RATE AND MEAN RESIDUAL FUNCTIONS

The failure rate function (FRF) $h(x)$ of the GTHPL distribution is given by

$$h(x) = \frac{f(x)}{1-F(x)} = \frac{\lambda e^{\lambda x}\left(\beta\left(e^{\lambda x}+\alpha-1\right)+2\eta\right)}{\left(e^{\lambda x}+\alpha-1\right)\left(\beta\left(e^{\lambda x}+\alpha-1\right)+\eta\right)}, \quad (13)$$

where $x > 0, \alpha, \eta, \lambda > 0, \beta \geq 0$.

Theorem 2.2 For all $\alpha, \eta, \lambda > 0, \beta \geq 0$, $h(x)$ is (i) decreasing if $0 < \alpha \leq \alpha_0$, where $1 < \alpha_0 < 1 + \frac{\eta}{\beta}$ such that $\beta^2\alpha_0{}^3 + (3\beta\eta - \beta^2)\alpha_0{}^2 - 2\eta^2 + 2\alpha_0\eta(\eta - 2\beta) = 0$.

(ii) unimodal if $\alpha_0 < \alpha < 1 + \frac{\eta}{\beta}$;

(iii) increasing if $\alpha \geq 1 + \frac{\eta}{\beta}$, with $h(0) = \frac{\lambda(\alpha\beta+2\eta)}{\alpha(\alpha\beta+\eta)}$ and $h(\infty) = \lambda$.

Proof. The first derivative of $h(x)$ is

$$h'(x) = \frac{\lambda^2 e^{\lambda x}}{\left(e^{\lambda x}+\alpha-1\right)^2\left(\beta\left(e^{\lambda x}+\alpha-1\right)+\eta\right)^2}\xi(t),$$

where $\xi(t) = \beta(\beta(\alpha-1)-\eta)t^2 + 2\beta(\alpha-1)(\beta(\alpha-1)+\eta)t + (\alpha-1)(\beta(\alpha-1)+2\eta)(\beta(\alpha-1)+\eta)$

and $e^{\lambda x} = t > 1$. Note that $\xi(t) > 0$ if $\alpha \geq 1 + \frac{\eta}{\beta}$ and $\xi(t) < 0$ if $\alpha \leq 1$. For $1 < \alpha \leq 1 + \frac{\eta}{\beta}$, there is a unique maximum at point $t_0 = \frac{(1-\alpha)(\beta(\alpha-1)+\eta)}{(\beta(\alpha-1)-\eta)}$. Since $\xi(\infty) = -\infty$ and $\xi(1) = \beta^2\alpha^3 + (3\beta\eta - \beta^2)\alpha^2 + 2\alpha\eta(\eta - 2\beta) - 2\eta^2$ $\xi(1) = \beta^2\alpha^3 + (3\beta\eta - \beta^2)\alpha^2 + 2\alpha\eta(\eta - 2\beta) - 2\eta^2$, it follows that $\xi(t) < 0$ if $\xi(1) < 0$. The equation $\beta^2\alpha_0{}^3 + (3\beta\eta - \beta^2)\alpha_0{}^2 + 2\alpha_0\eta(\eta - 2\beta) - 2\eta^2 = 0$ has a unique zero point. The proof is complete.

Moreover, the mean residual life function (MRLF) of the GTHPL distribution is defined as follows:

$$\begin{aligned} \mu(x) &= E(X - x \mid X > x) \\ &= \frac{1}{1-F(x)}\int_x^\infty 1 - F(t)dt \\ &= \frac{\left(e^{\lambda x}+\alpha-1\right)^2}{\lambda(1-\alpha)\left(\beta\left(e^{\lambda x}+\alpha-1\right)+\eta\right)}\left(\frac{(\alpha-1)\beta+\eta}{\alpha-1}\ln\frac{e^{\lambda x}}{e^{\lambda x}+\alpha-1} + \frac{\eta}{e^{\lambda x}+\alpha-1}\right), x > 0. \end{aligned} \quad (14)$$

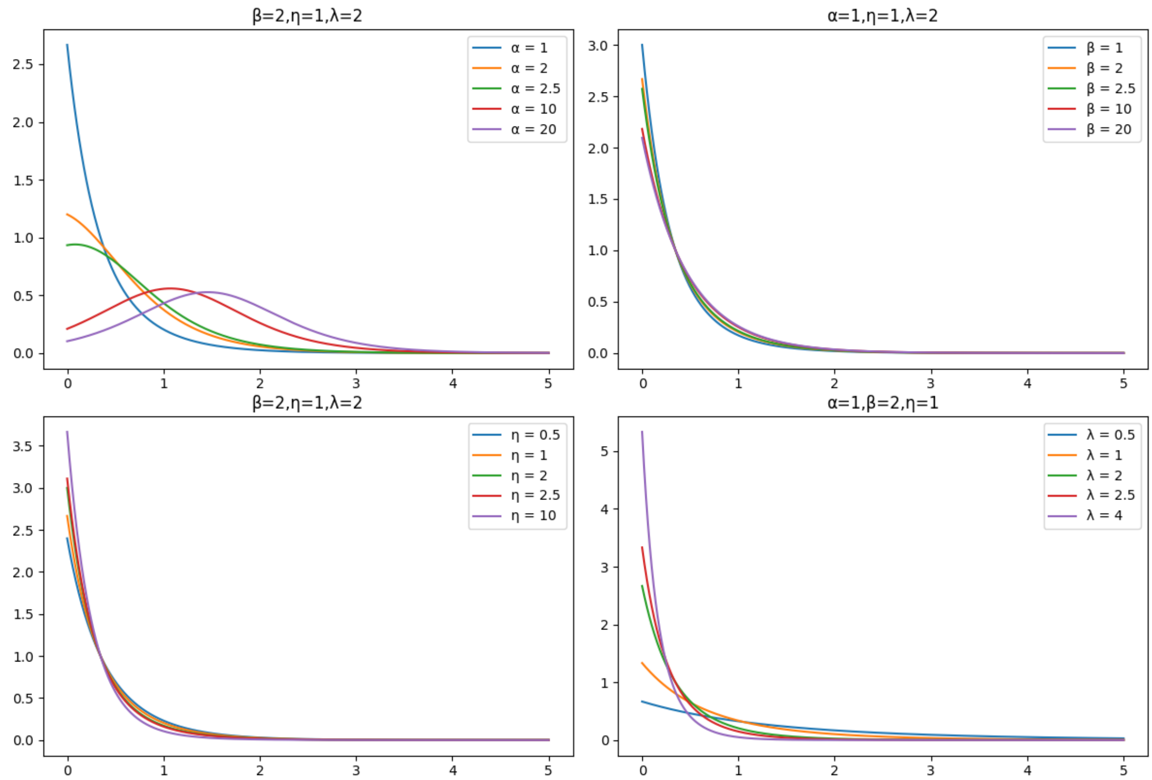

FIGURE 1. The pdf of the GTHPL distribution for various values of α, β, η and λ

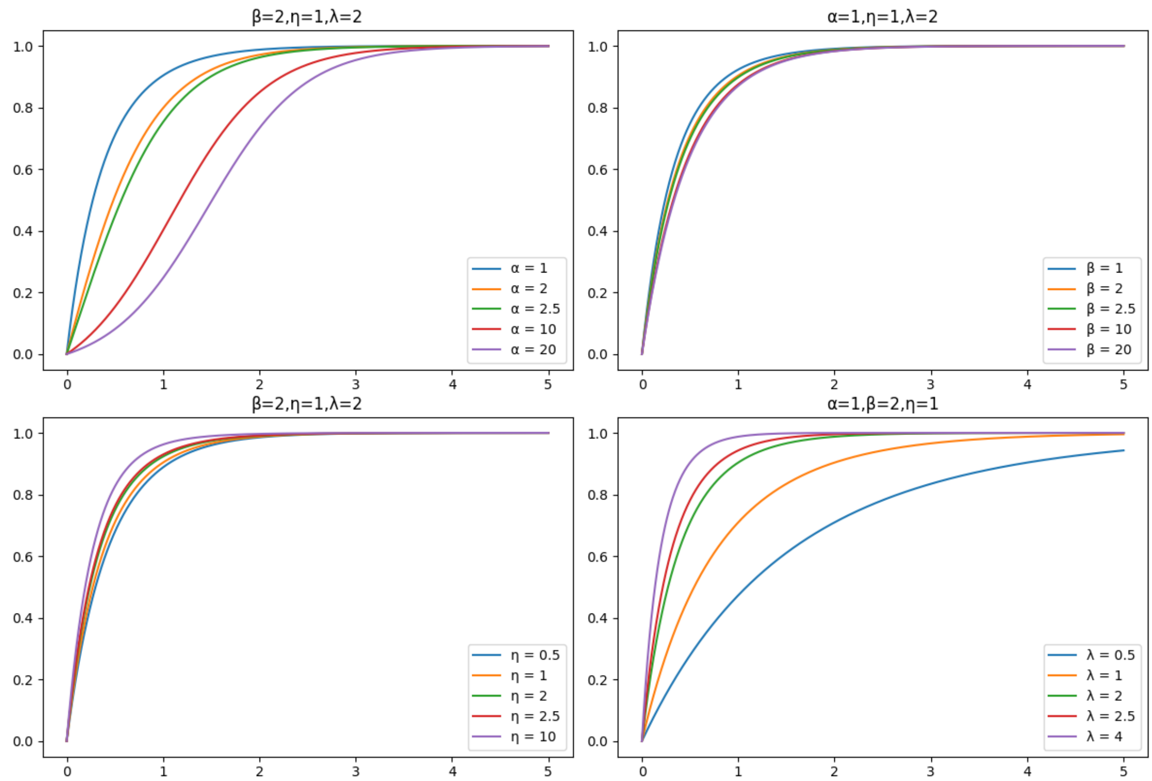

FIGURE 2. The cdf of the GTHPL distribution for various values of α, β, η and λ

LIMIT BEHAVIOR

Proposition 2.1 The asymptotic CDF, PDF, and FRF of the GTHPL distribution as $x \to 0$ are given by

$$1 - F(x) \sim \frac{\alpha^2}{\alpha\beta+\eta}\frac{(\beta(\lambda x+\alpha)+\eta)}{(\lambda x+\alpha)^2}, \tag{15}$$

$$f(x) \sim \frac{\lambda\alpha^2}{\alpha+\beta}\frac{(\beta(\lambda x+\alpha)+2\eta)}{(\lambda x+\alpha)^3}, \tag{16}$$

$$h(x) \sim \frac{\lambda(\beta(\lambda x+\alpha)+2\eta)}{(\lambda x+\alpha)(\beta(\lambda x+\alpha)+\eta)}. \tag{17}$$

Proof According to the MacLaurin series, when $x \to 0$, we have $e^{\lambda x} \to 1$ and $e^{\lambda x} - 1 \sim \lambda x e^{\lambda x} - 1 \sim \lambda x$. The proof is complete.

Proposition 2.2 The asymptotic CDF, PDF, and FRF of the GTHPL distribution as $x \to \infty$ are given by

$$1 - F(x) \sim \frac{\alpha^2\beta}{\alpha\beta+\eta}e^{-\lambda x}, \tag{18}$$

$$f(x) \sim \frac{\lambda\alpha^2\beta}{\alpha\beta+\eta}e^{-\lambda x}, \tag{19}$$

$$h(x) \sim \lambda. \tag{20}$$

Proof According to the MacLaurin series, when $x \to \infty$, we have $e^{\lambda x} + c \to e^{\lambda x}$, where c is a constant. The proof is complete.

MOMENTS AND ASSOCIATED MEASURES

We derive the rth raw moment (about the origin) of the $\mathrm{GTHPL}(\alpha,\beta,\eta,\lambda)$ distribution. We consider the following two cases:

(i) $\alpha \neq 1$: For positive integer r, we have

$$\begin{aligned}\mu_r = E(x^r) &= \frac{\alpha^2 r}{\alpha\beta+\eta}\int_0^\infty \frac{(\beta(e^{\lambda x}+\alpha-1)+\eta)x^{r-1}}{(e^{\lambda x}+\alpha-1)^2}dx \\ &= \frac{\alpha^2\Gamma(r+1)}{\lambda^r\bar{\alpha}^2(\alpha\beta+\eta)}\left((\bar{\alpha}\beta-\eta)L_r(\bar{\alpha})+\eta L_{r-1}(\bar{\alpha})\right),\end{aligned} \tag{21}$$

where $\bar{\alpha} = 1 - \alpha$ and

$$L_s(z) = \frac{z}{\Gamma(s)}\int_0^\infty \frac{t^{s-1}}{e^t - z}\mathrm{d}t, s > 0, z < 1, \tag{22}$$

is the polynomial function, where $L_1(z) = -\ln(1-z)$ and $L_0(z) = z\frac{\partial L_1(z)}{\partial z} = \frac{z}{1-z}$.

Moreover, the polynomial function satisfies

$$L_{s-1}(z) - L_s(z) = \frac{z^2}{\Gamma(s)}\int_0^\infty \frac{t^{s-1}}{(e^t - z)^2}\mathrm{d}t, \tag{23}$$

$$L_{s-1}(z) - 3L_s(z) + 2L_{s+1}(z) = \frac{2z^3}{\Gamma(s+1)}\int_0^\infty \frac{t^s}{(e^t - z)^3}\mathrm{d}t. \tag{24}$$

(ii) $\alpha = 1$, For all positive integers r,

$$\begin{aligned}\mu_r = E(X^r) &= \frac{1}{\beta+\eta}\left(\int_0^\infty \frac{\lambda x^r\beta}{e^{\lambda x}}dx + \int_0^\infty \frac{2\eta\lambda x^r e^{\lambda x}}{e^{3\lambda x}}dx\right) \\ &= \frac{r!(2^r\beta+\eta)}{(2\lambda)^r(\beta+\eta)}.\end{aligned} \tag{25}$$

Hence, the first four raw moments are given as follows:

$$\mu = \begin{cases}\frac{\alpha((\eta-\bar{\alpha}\beta)\alpha\ln\alpha+\eta\bar{\alpha})}{\lambda\bar{\alpha}^2(\alpha\beta+\eta)}, & \alpha \neq 1,\\ \frac{2\beta+\eta}{2\lambda(\beta+\eta)}, & \alpha = 1,\end{cases} \tag{26}$$

$$\mu_2 = \begin{cases}\frac{2\alpha^2((\bar{\alpha}\beta-\eta)L_2(\bar{\alpha})-\eta\ln\alpha)}{\lambda^2\bar{\alpha}^2(\alpha\beta+\eta)}, & \alpha \neq 1,\\ \frac{4\beta+\eta}{2\lambda^2(\beta+\eta)}, & \alpha = 1,\end{cases} \tag{27}$$

$$\mu_3 = \begin{cases}\frac{6\alpha^2((\bar{\alpha}\beta-\eta)L_3(\bar{\alpha})+\eta L_2(\bar{\alpha}))}{\lambda^3\bar{\alpha}^2(\alpha\beta+\eta)}, & \alpha \neq 1,\\ \frac{24\beta+3\eta}{4\lambda^3(\beta+\eta)}, & \alpha = 1,\end{cases} \tag{28}$$

$$\mu_4 = \begin{cases}\frac{24\alpha^2((\bar{\alpha}\beta-\eta)L_4(\bar{\alpha})+\eta L_3(\bar{\alpha}))}{\lambda^4\bar{\alpha}^2(\alpha\beta+\eta)}, & \alpha \neq 1,\\ \frac{48\beta+3\eta}{2\lambda^4(\beta+\eta)}, & \alpha = 1.\end{cases} \tag{29}$$

Figure 3 displays the first four moments of the GTHPL distribution for various values of α, β, η and λ.

Therefore, the variance of the GTHPL distribution is

$$\sigma^2 = \begin{cases}\frac{-\alpha^2(2\bar{\alpha}^2(\alpha\beta+\eta)((\eta-\bar{\alpha}\beta)L_2(\bar{\alpha})+\eta\ln\alpha)+((\eta-\bar{\alpha}\beta)\alpha\ln\alpha+\eta\bar{\alpha})^2)}{\lambda^2\bar{\alpha}^4(\alpha\beta+\eta)^2}, & \alpha \neq 1,\\ \frac{\eta^2+6\beta\eta+4\beta^2}{4(\lambda\beta+\lambda\eta)^2}, & \alpha = 1.\end{cases} \tag{30}$$

Moreover, the skewness and kurtosis of the GTHPL distribution can be given as follows:

$$\text{Skewness} = \frac{E(X-\mu)^3}{\sigma^3} = \frac{\mu_3 - 3\mu_2\mu + 2\mu^3}{\sigma^3}, \tag{31}$$

$$\text{Kurtosis} = \frac{E(X-\mu)^4}{\sigma^4} = \frac{\mu_4 - 4\mu_3\mu + 6\mu_2\mu^2 - 3\mu^4}{\sigma^4}. \tag{32}$$

The moment generation function of the GTHPL distribution via the exponential function of Maclaurin's series expansion is

$$M_X(t) = E(e^{tX}) = \sum_{r=0}^{\infty}\frac{t^r}{r!}E(X^r). \tag{33}$$

For $\alpha \neq 1$, we have

$$M_X(t) = \frac{\alpha^2}{\bar{\alpha}^2(\alpha\beta+\eta)}\sum_{r=1}^{\infty}\frac{t^r\Gamma(r+1)((\bar{\alpha}\beta-\eta)L_r(\bar{\alpha})+\eta L_{r-1}(\bar{\alpha}))}{\lambda^r r!}. \tag{34}$$

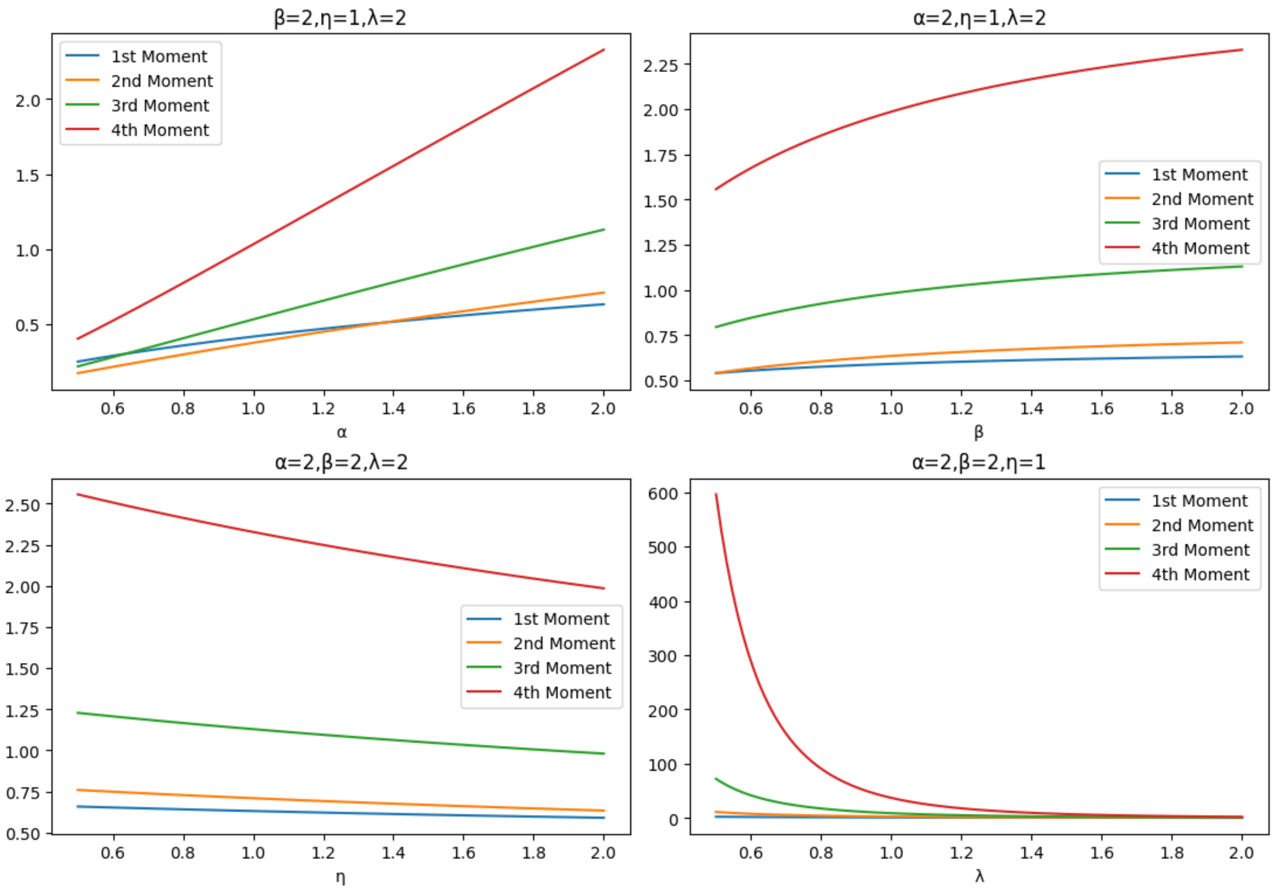

FIGURE 3. The first four moments of the GTHPL distribution for various values of α, β, η and λ

For $\alpha = 1$, we have

$$M_X(t) = E(e^{tX}) = \sum_{r=0}^{\infty} \frac{t^r(2^r\beta+\eta)}{(2\lambda)^r(\beta+\eta)}. \qquad (35)$$

STOCHASTIC ORDERS

Numerous stochastic orders exist and have various applications (Shaked & Shanthikumar 2007). In this context, the likelihood ratio order $\leq_{\mathrm{lr}}$, the usual stochastic order $\leq_{\mathrm{st}}$, the failure rate order $\leq_{\mathrm{fr}}$, and the mean residual life order $\leq_{\mathrm{mrl}}$ are considered. A random variable X is considered to be less than a random variable Y:

(i) if $F_X(x) \geq F_Y(x), X \leq_{\mathrm{st}} Y$ for all x;
(ii) if $h_X(x) \geq h_Y(x),\ X \leq_{\mathrm{fr}} Y$ for all x;
(iii) if $\mu_X(x) \leq \mu_Y(x), X \leq_{\mathrm{mrl}} Y$ for all x;
(iv) if $\frac{f_X(x)}{f_Y(x)},\ X \leq_{\mathrm{lr}} Y$ decreases in x.

Theorem 2.3 Let $X \sim \mathrm{GTHPL}(\alpha_1, \beta, \eta, \lambda)$ and $Y \sim \mathrm{GTHPL}(\alpha_2, \beta, \eta, \lambda)$. For all $\eta, \lambda > 0, \beta \geq 0$ and $0 < \alpha_1 \leq \alpha_2, X \leq_{\mathrm{lr}} Y$ (and hence $X \leq_{\mathrm{fr}} Y, X \leq_{\mathrm{mrl}} Y$ and $X \leq_{\mathrm{st}} Y$).

Proof For $\beta = 0$, we have

$$\frac{d}{dx} \ln \frac{f_X(x)}{f_Y(x)} = \frac{3\lambda e^{\lambda x}(\alpha_1 - \alpha_2)}{(e^{\lambda x} + \alpha_1 - 1)(e^{\lambda x} + \alpha_2 - 1)}.$$

For $\beta > 0$, letting $\gamma = 2\eta/\beta$, we have

$$\frac{d}{dx} \ln \frac{f_X(x)}{f_Y(x)} = \lambda e^{\lambda x}(\alpha_2 - \alpha_1)\left(\frac{1}{(e^{\lambda x} + \alpha_1 - 1 + \gamma)(e^{\lambda x} + \alpha_2 - 1 + \gamma)} - \frac{3}{(e^{\lambda x} + \alpha_1 - 1)(e^{\lambda x} + \alpha_2 - 1)}\right).$$

Since $0 < \alpha_1 \leq \alpha_2,\ \gamma, \lambda > 0$, and hence $\frac{d}{dx} \ln \frac{f_X(x)}{f_Y(x)} \leq 0$, it follows that $\frac{f_X(x)}{f_Y(x)}$ is decreasing in x. Then $X \leq_{\mathrm{lr}} Y$.

Note that $X \leq_{\mathrm{lr}} Y$ implies that $X \leq_{\mathrm{fr}} Y,\ X \leq_{\mathrm{mrl}} Y$ and that $X \leq_{st} Y$. The proof is complete.

ESTIMATION THEORY

MAXIMUM LIKELIHOOD ESTIMATION

Consider $x_1, x_2, \ldots, x_n$ as a random sample from the GTHPL distribution with parameters $\Theta = (\alpha, \beta, \eta, \lambda)$. The loglikelihood function is given as follows:

$$l(\Theta) = n\ln\frac{\lambda\alpha^2}{(\alpha\beta+\eta)} + \lambda\sum_{i=1}^{n} x_i + \sum_{i=1}^{n} \ln\left(\beta\left(e^{\lambda x_i} + \alpha - 1\right) + 2\eta\right) - 3\sum_{i=1}^{n} \ln\left(e^{\lambda x_i} + \alpha - 1\right). \quad (36)$$

Hence, the maximum likelihood estimators (MLEs) $\widehat{\Theta} = (\hat{\alpha}, \hat{\beta}, \hat{\eta}, \hat{\lambda})$ of the parameters $\Theta = (\alpha, \beta, \eta, \lambda)$ are the simultaneous solutions of the following equations:

$$\begin{cases} \frac{\partial l(\Theta)}{\partial \alpha} = \frac{2n}{\alpha} + \sum_{i=1}^{n} \frac{\beta}{\beta(e^{\lambda x_i}+\alpha-1)+2\eta} - \frac{n\beta}{\alpha\beta+\eta} - 3\sum_{i=1}^{n} \frac{1}{e^{\lambda x_i}+\alpha-1} = 0 \\ \frac{\partial l(\Theta)}{\partial \beta} = \sum_{i=1}^{n} \frac{e^{\lambda x_i}+\alpha-1}{\beta(e^{\lambda x_i}+\alpha-1)+2\eta} - \frac{n\alpha}{\alpha\beta+\eta} = 0 \\ \frac{\partial l(\Theta)}{\partial \eta} = 2\sum_{i=1}^{n} \frac{1}{\beta(e^{\lambda x_i}+\alpha-1)+2\eta} - \frac{n}{\alpha\beta+\eta} = 0 \\ \frac{\partial l(\Theta)}{\partial \lambda} = \frac{n}{\lambda} + \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} \frac{\beta x_i e^{\lambda x_i}}{\beta(e^{\lambda x_i}+\alpha-1)+2\eta} - 3\sum_{i=1}^{n} \frac{x_i e^{\lambda x_i}}{e^{\lambda x_i}+\alpha-1} = 0. \end{cases} \quad (37)$$

Since the maximum likelihood equations (37) are nonlinear in nature, the solutions $\widehat{\Theta} = (\hat{\alpha}, \hat{\beta}, \hat{\eta}, \hat{\lambda})$ have no closed form. However, these equations can be solved numerically via numerical optimization algorithms. In this work, the unknown parameters $(\alpha, \beta, \eta, \lambda)$ are estimated via maximization via the differential evolution function of Python software. Compared with the Newton–Raphson or quasi-Newton–Raphson methods, which rely on gradient information (e.g., first or second derivatives) from the log-likelihood function, parameter estimation via differential evolution algorithms does not require explicit solutions for derivatives, which renders differential evolution suitable for problems in which computational access to gradient information is difficult or nonexistent.

SIMULATION STUDY

In this section, we conduct a simulation study to generate random variables from the GTHPL model. The suitability of the maximum likelihood estimation (MLE) method for estimating unknown model parameters is investigated. To evaluate the accuracy and consistency of the estimates, we calculated the bias and mean squared error (MSE) of the MLE of the parameters. The calculations pertaining to the study were carried out via Python software, version 3.10, with the help of self-programmed codes. Moreover, the scipy.optimise package was used to obtain the maximum likelihood estimates of the parameters from GTHPL.

Using the inversion method, we can generate random numbers from the GTHPL distribution via the following equation

$$x_i = F^{-1}(u_{(i)}) = \lambda^{-1}\ln\left(\frac{\alpha\left(\left(\alpha^2\beta^2+4(1-u_{(i)})(\alpha\beta+\eta)\eta\right)^{\frac{1}{2}}+\alpha\beta\right)}{2(\alpha\beta+\eta)(1-u_{(i)})} - \alpha + 1\right), \quad (38)$$

where $u_{(i)}$ is a uniformly distributed random variable, $U(0,1)$. Given the sample size n, for each $u_{(i)}$, $i = 1, 2, \ldots, n$, we can solve the above system of equations for x_i $(i = 1, 2, \ldots, n)$ simultaneously. Hence, one can generate random numbers when α, β, η and λ are known. Since $\alpha, \eta, \lambda > 0$ and $\beta \geq 0$. In the simulation study, we select $\alpha = 1.0$, $\eta = 4.0$, $\lambda = 2.0$ and $\beta = 0.0$, 1.0 and 2.0 for sample sizes $n = 20, 50, 100$ and 300, respectively.

In each simulation, for a given parameter combination $\Theta = (\alpha, \beta, \eta, \lambda)$, we first resample the observations from the GTHPL distribution N=10,000 times to obtain the observed values $(x_1, x_2, \ldots, x_n)$. Using differential evolution algorithms, we separately calculate the average estimates (AEs), average bias (ABs), and average mean squared error (AMSE) as follows:

(i) The average estimates (AEs) of the MLEs $\widehat{\Theta} = (\hat{\alpha}, \hat{\beta}, \hat{\eta}, \hat{\lambda})$ are given by

$$AE(\widehat{\Theta}) = \frac{1}{N}\sum_{i=1}^{N} \widehat{\Theta}_i, i = 1, 2, 3, 4,$$

(ii) The average biases (ABs) of the MLEs $\widehat{\Theta} = (\hat{\alpha}, \hat{\beta}, \hat{\eta}, \hat{\lambda})$ are given by

$$AB(\widehat{\Theta}) = \frac{1}{N}\sum_{i=1}^{N} (\widehat{\Theta}_i - \Theta_i), i = 1, 2, 3, 4.$$

(iii) The average mean square errors (AMSEs) of the MLEs $\widehat{\Theta} = (\hat{\alpha}, \hat{\beta}, \hat{\eta}, \hat{\lambda})$ are given by

$$AMSE(\widehat{\Theta}) = \frac{1}{N}\sum_{i=1}^{N} (\widehat{\Theta}_i - \Theta_i)^2, i = 1, 2, 3, 4.$$

The simulation results are summarized in Table 1.

As shown in Table 1, when the sample size n increases, the estimated values of $(\alpha, \beta, \eta, \lambda)$ obtained through the MLE method converge to the true parameter values. Simultaneously, the biases and mean squared errors for each parameter decrease and gradually approach zero. Hence, the simulation results illustrate the consistency property of the MLEs $(\hat{\alpha}, \hat{\beta}, \hat{\eta}, \hat{\lambda})$.

THE PROPOSED GTHPL REGRESSION MODEL

GTHPL REGRESSION AND MLES

In this section, a new regression model based on the GTHPL distribution is presented. According to Equation (26), we parameterize

$$\lambda = \begin{cases} \frac{\alpha((\eta-\bar{\alpha}\beta)\alpha\ln\alpha+\eta\bar{\alpha})}{\mu\bar{\alpha}^2(\alpha\beta+\eta)}, & \alpha \neq 1, \\ \frac{2\beta+\eta}{2\mu(\beta+\eta)}, & \alpha = 1, \end{cases} \tag{39}$$

to the pdf of the GTHPL distribution, then we have

$$p(y \mid \alpha,\beta,\eta,\mu) =$$

$$p(y \mid \alpha,\beta,\eta,\mu) = \begin{cases} \frac{\alpha^3 e^{\frac{\alpha((\eta-\bar{\alpha}\beta)\alpha\ln\alpha+\eta\bar{\alpha})}{\mu\bar{\alpha}^2(\alpha\beta+\eta)}y}\left(\beta\left(e^{\frac{\alpha((\eta-\bar{\alpha}\beta)\alpha\ln\alpha+\eta\bar{\alpha})}{\mu\bar{\alpha}^2(\alpha\beta+\eta)}y}+\alpha-1\right)+2\eta\right)}{\mu\bar{\alpha}^2(\alpha\beta+\eta)^2\left(e^{\frac{\alpha((\eta-\bar{\alpha}\beta)\alpha\ln\alpha+\eta\bar{\alpha})}{\mu\bar{\alpha}^2(\alpha\beta+\eta)}y}+\alpha-1\right)^3((\eta-\bar{\alpha}\beta)\alpha\ln\alpha+\eta\bar{\alpha})^{-1}}, & \alpha \neq 1, \\ \frac{\alpha^2(2\beta+\eta)e^{\frac{2\beta+\eta}{2\mu(\beta+\eta)}y}\left(\beta\left(e^{\frac{2\beta+\eta}{2\mu(\beta+\eta)}y}+\alpha-1\right)+2\eta\right)}{2\mu(\beta+\eta)(\alpha\beta+\eta)\left(e^{\frac{2\beta+\eta}{2\mu(\beta+\eta)}y}+\alpha-1\right)^3}, & \alpha = 1, \end{cases} \tag{40}$$

where $\alpha,\eta > 0, \alpha\beta+\eta+\beta > 0$ and $E(Y \mid \alpha,\beta,\eta,\mu) = \mu$. The below log-link function is used to link the explanatory variable to the mean of the response variable. It can be expressed as

$$\log(\mu^{(i)}) = x_i^T\gamma, i = 1,\dots,n, \tag{41}$$

where $x_i^T = (1, x_{i1}, x_{i2}, \dots, x_{ik})$ is the vector of covariates and $\gamma = (\gamma_0, \gamma_1, \gamma_2, \dots, \gamma_k)^T$ is the unknown vector of regression coefficients.

The log-likelihood function of the GTHPL regression model is as follows:

(1) For $\alpha \neq 1$, we have

$$\ell(\alpha,\beta,\eta,\gamma) = 3n\ln\alpha - 2n\ln\bar{\alpha} + \sum_{i=1}^{n}\frac{\alpha((\eta-\bar{\alpha}\beta)\alpha\ln\alpha+\eta\bar{\alpha})}{exp(x_i^T\gamma)\bar{\alpha}^2(\alpha\beta+\eta)}y_i - 2n\ln(\alpha\beta+\eta) - \sum_{i=1}^{n}x_i^T\gamma$$
$$+\sum_{i=1}^{n}\ln\left(\beta\left(e^{\frac{\alpha((\eta-\bar{\alpha}\beta)\alpha\ln\alpha+\eta\bar{\alpha})}{exp(x_i^T\gamma)\bar{\alpha}^2(\alpha\beta+\eta)}y_i}+\alpha-1\right)+2\eta\right)$$
$$+n\ln((\eta-\bar{\alpha}\beta)\alpha\ln\alpha+\eta\bar{\alpha})$$
$$-3\sum_{i=n}^{n}\left(e^{\frac{\alpha((\eta-\bar{\alpha}\beta)\alpha\ln\alpha+\eta\bar{\alpha})}{exp(x_i^T\gamma)\bar{\alpha}^2(\alpha\beta+\eta)}y_i}+\alpha-1\right). \tag{42}$$

(2) For $\alpha = 1$, we have

$$\ell(\alpha,\beta,\eta,\gamma) = 2n\ln\alpha +$$
$$\sum_{i=1}^{n}\ln\left(\beta\left(e^{\frac{2\beta+\eta}{2\exp(x_i^T\gamma)(\beta+\eta)}y_i}+\alpha-1\right)+2\eta\right)$$
$$+\text{nln}\frac{(2\beta+\eta)}{(\alpha\beta+\eta)}+\sum_{i=1}^{n}\frac{2\beta+\eta}{2\,exp(x_i^T\gamma)(\beta+\eta)}y_i$$
$$-3\sum_{i=1}^{n}\ln\left(e^{\frac{2\beta+\eta}{2\exp(x_i^T\gamma)(\beta+\eta)}y_i}+\alpha-1\right)-\sum_{i=1}^{n}x_i^T\gamma$$
$$-\text{nln}\,2 - nln\frac{2\beta+\eta}{2\,exp(x_i^T\gamma)(\beta+\eta)}. \tag{43}$$

The MLEs of $(\alpha,\beta,\eta,\gamma)$ are obtained via direct maximization of the log-likelihood function. In this paper, the minimize function of Python software is applied to calculate the MLEs $(\hat{\alpha},\hat{\beta},\hat{\eta},\hat{\gamma})$.

SIMULATION STUDY

Similar to simulation study in Estimation Theory section, we conducted a simulation study to evaluate the performance of the MLEs of unknown parameters in the GTHPL regression model.

Suppose that the log-link function in Equation (41) can be expressed as

$$\log(\mu^{(i)}) = \gamma_0 + \gamma_1 x_{i1} + \gamma_2 x_{i2}, i = 1,\dots,n, \tag{44}$$

where x_{i1} and x_{i2} are generated from U (0,1). Here, the sample sizes are $n = 20, 50, 100$ and 300. Moreover, owing the constraint $\beta \geq 0$, $\alpha,\eta > 0$. Using the log-likelihood functions (42) and (43) of the GTHPL regression model, the following parameter settings are used.

$\alpha = 1.0$, $\beta = 2.0$, $\eta = 4.0$, $\gamma_0 = 1.0$, $\gamma_1 = 0.5$, $\gamma_2 = 0.5$;

$\alpha = 1.0$, $\beta = 2.0$, $\eta = 6.0$, $\gamma_0 = 1.0$, $\gamma_1 = -2.0$, $\gamma_2 = 0.5$;

$\alpha = 1.5$, $\beta = 2.0$, $\eta = 4.0$, $\gamma_0 = 1.0$, $\gamma_1 = 0.5$, $\gamma_2 = 0.5$;

$\alpha = 1.5$, $\beta = 2.0$, $\eta = 6.0$, $\gamma_0 = 1.0$, $\gamma_1 = -2.0$, $\gamma_2 = 0.5$;

In each simulation, for a given parameter combination $\Theta' = (\alpha,\beta,\eta,\gamma_0,\gamma_1,\gamma_2)$, we first resample the observations from the GTHPL linear model N=10,000 times to obtain the observed values $(x_1, x_2, \dots, x_n)$ and $(y_1, y_2, \dots, y_n)$. Using differential evolution algorithms, we separately calculate the AEs, ABs, and AMSEs

TABLE 1. Maximum likelihood estimates of the GTHPL distribution

Sample size	Parameters	$\alpha = 1.0$	$\beta = 0.0$	$\eta = 4.0$	$\lambda = 2.0$
		$\hat{\alpha}$	$\hat{\beta}$	$\hat{\eta}$	$\hat{\lambda}$
n=20	AE	1.5413	-0.2084	3.6234	3.2977
	AB	0.5413	-0. 2084	-0.3766	1.2977
	AMSE	0.8775	0.0440	0.1563	1.8275
n=50	AE	1.3488	-0.1832	3.6172	3.0734
	AB	0.3488	-0.1832	-0.3828	1.7035
	AMSE	0.3465	0.0369	0.1548	1.5602
n=100	AE	1.1670	-0.1381	3.7503	2.8494
	AB	0.1670	-0.1381	-0.2497	0.7494
	AMSE	0.2958	0.0325	0.1477	1.0365
n=300	AE	1.0670	-0.0611	3.9155	2.2716
	AB	0.0860	-0.0611	-0.0845	0.2716
	AMSE	0.1834	0.0268	0.0930	0.6589
Sample size	Parameters	$\alpha = 1.0$	$\beta = 1.0$	$\eta = 4.0$	$\lambda = 2.0$
		$\hat{\alpha}$	$\hat{\beta}$	$\hat{\eta}$	$\hat{\lambda}$
n=20	AE	1.3689	1.4296	4.6691	2.3019
	AB	0.3689	0.4296	0.6191	0.3019
	AMSE	0.7835	1.0779	1.6503	1.3331
n=50	AE	1.2026	1.3742	4.4891	2.2438
	AB	0.2026	0.3742	0.4191	0.2438
	AMSE	0.4277	1.1973	1.3503	0.9278
n=100	AE	1.1116	1.1806	3.7903	2.2184
	AB	0.1116	0.1806	-0.2097	0.2184
	AMSE	0.2157	1.2044	0.3326	0.6400
n=300	AE	1.0354	1.0306	3.9948	2.0399
	AB	0.0354	0.0706	-0.0052	0.0399
	AMSE	0.0840	0.0944	0.0418	0.2577
Sample size	Parameters	$\alpha = 1.0$	$\beta = 2.0$	$\eta = 4.0$	$\lambda = 2.0$
		$\hat{\alpha}$	$\hat{\beta}$	$\hat{\eta}$	$\hat{\lambda}$
n=20	AE	0.9592	1.2990	4.1708	1.9931
	AB	-0.0408	-0.7010	0.1708	-0.0069
	AMSE	0.1188	2.1957	0.1222	0.2388
n=50	AE	0.9101	1.7313	3.9934	1.7153
	AB	-0.0899	-0.2687	-0.0066	-0.2847
	AMSE	0.1646	1.6312	0.0841	0.4557
n=100	AE	1.1642	0.9495	4.1781	2.0684
	AB	0.1642	-1.0505	0.1781	0.0684
	AMSE	0.2107	2.0583	0.1108	0.3898
n=300	AE	1.0968	1.9878	4.0070	2.1864
	AB	0.0968	-0.0122	0.0070	0.1864
	AMSE	0.0350	0.1949	0.0723	0.1021

of the MLEs $\widehat{\Theta}' = (\hat{\alpha}, \hat{\beta}, \hat{\eta}, \widehat{\gamma_0}, \widehat{\gamma_1}, \widehat{\gamma_2})$. The simulation results are summarized in Table 2.

The data presented in Table 2 indicates that with increasing sample size (n), the maximum likelihood estimates of $(\alpha, \beta, \eta, \gamma_0, \gamma_1, \gamma_2)$ converge asymptotically to the true parameter values. In conjunction with this convergence, the observed biases and mean squared errors for the respective parameters progressively diminish, ultimately approaching zero. These findings from the simulation study provide empirical evidence for the consistency of the maximum likelihood estimators $(\hat{\alpha}, \hat{\beta}, \hat{\eta}, \widehat{\gamma_0}, \widehat{\gamma_1}, \widehat{\gamma_2})$.

REAL DATA ANALYSIS

APPLICATION OF THE GTHPL MODEL

In this section, a real dataset is analyzed to demonstrate the adaptability of the GTHPL model. This dataset (Shen et al. 2018) consists of survival times for 217 female patients diagnosed with hepatocellular carcinoma, measured in years. Table 3 summarizes the descriptive statistical data of the dataset, and its histogram and boxplot are shown in Figure 4. From Table 3, the skewness value (1.3610) of the survival dataset indicates that this dataset is skewed to the left. For model comparison, this dataset was also applied to evaluate several well-known models, such as the log-normal, Gompertz, Weibull, Gompertz–Lindley (GL) and Gompertz–two–parameter–Lindley (GTPL) distributions.

Table 4 reports the MLEs, 95% confidence intervals and *p*-values of α, β, η and λ. Given a significance level of 0.05, the *p*-values of the tests for all the parameters α, β, η and λ in the GTHPL model are less than 0.05, and are therefore statistically significant. Moreover, Table 5 summarizes the Akaike information criterion (AIC), Bayesian information criterion (BIC), consistent Akaike information criterion (CAIC) and Hannan–Quinn information criterion (HQIC) of the six models. The GTHPL model has lower AIC, CAIC, BIC, and HQIC values than the log-normal, Weibull, Gompertz, GL and GTPL models, which indicates that the GTHPL distribution performs better than the other five models. Hence, the proposed distribution provides a better fit to this real survival dataset.

APPLICATION OF THE GTHPL REGRESSION MODEL

As an extension of the GTHPL model, we built a GTHPL regression model for the dataset (Shen et al. 2018) mentioned earlier to assess whether the proposed regression model would provide a better fit than the GL and GTPL regression models for this dataset. After data preprocessing, such as outliers, missing values and normalization, the remainder of this dataset consists of survival times for 185 female patients diagnosed with hepatocellular carcinoma y_i (i=1, 2, ..., 185) with the following two covariates: x_{i1} represents the aspartate transaminase level (U/L). x_{i2} represents the total bilirubin concentration (μmol/L). The simulation results of the GL, GTPL and GTHPL regression models are shown in Table 6.

Table 6 shows that when the values of aspartate aminotransferase (U/L) and total bilirubin (μmol/L) increase, the survival time of female patients also increases. These findings indicate that these two covariates have a positive effect on the survival time of female patients. Given a significance level of 0.05, the *p*-values of the tests for all the parameters in the GL, GTPL, and GTHPL regression models are less than 0.05 and are therefore statistically significant. Moreover, for better comparison, Table 7 summarizes the AIC, BIC, CAIC and HQIC of the three regression models. Obviously, the GTHPL regression model has lower AIC, CAIC, BIC, and HQIC values than the GL regression model and the GTPL regression model does, which implies that the new model has a better ability to fit the data. Finally, compared with the six models (log-normal, Gompertz, Weibull, GL, GTPL and GTHPL) without covariates in Table 5, the GTHPL regression model has the best data fitting effect. Therefore, as an alternative, we recommend the use of GTHPL regression model to fit this dataset.

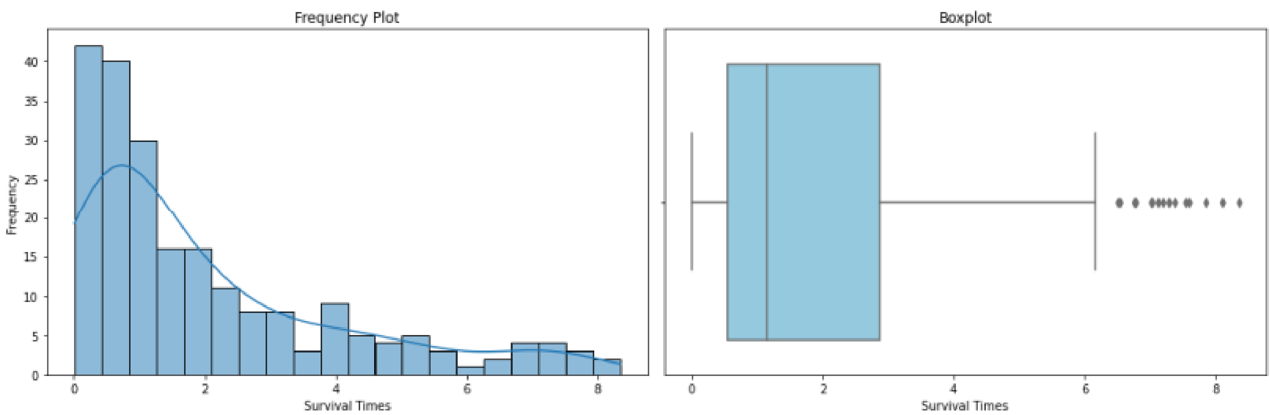

FIGURE 4. Histogram and box plot for the survival data

TABLE 2. Maximum likelihood estimates of the GTHPL regression model

Sample size	Parameters	$\alpha = 1.0$	$\beta = 2.0$	$\eta = 4.0$	$\gamma_0 = 1.0$	$\gamma_1 = 0.5$	$\gamma_2 = 0.5$
		$\hat{\alpha}$	$\hat{\beta}$	$\hat{\eta}$	$\widehat{\gamma_0}$	$\widehat{\gamma_1}$	$\widehat{\gamma_2}$
n=20	AE	0.1776	1.3270	3.5163	0.9268	0.4642	0.4966
	AB	-0.8224	-0.6730	-0.4837	-0.0732	-0.0358	-0.0034
	AMSE	0.6763	0.4529	0.2339	0.0054	0.0024	0.0077
n=50	AE	0.5333	1.4454	3.6268	0.9878	0.4753	0.5018
	AB	-0.4667	-0.5546	-0.3732	-0.0122	-0.0247	0.0018
	AMSE	0.2178	0.3075	0.1393	0.0011	0.0008	0.0001
n=100	AE	0.8766	1.5911	3.7526	0.9916	0.4974	0.4949
	AB	-0.1234	-0.4089	-0.2474	-0.0084	-0.0026	-0.0051
	AMSE	0.0152	0.1672	0.0612	0.0002	0.0000	0.0003
n=300	AE	1.0763	1.9739	4.0204	0.9952	0.4975	0.4980
	AB	0.0763	-0.0261	0.0204	-0.0048	-0.0025	-0.0020
	AMSE	0.0058	0.0007	0.0004	0.0000	0.0000	0.0000
Sample size	Parameters	$\alpha = 1.0$	$\beta = 2.0$	$\eta = 6.0$	$\gamma_0 = 1.0$	$\gamma_1 = -2.0$	$\gamma_2 = 0.5$
		$\hat{\alpha}$	$\hat{\beta}$	$\hat{\eta}$	$\widehat{\gamma_0}$	$\widehat{\gamma_1}$	$\widehat{\gamma_2}$
n=20	AE	0.8916	2.2499	6.0982	0.5210	-1.6753	0.3267
	AB	-0.1084	0.2499	0.0982	-0.4790	0.3247	-0.1733
	AMSE	0.0117	0.0624	0.0097	2.2682	1.0513	0.2940
n=50	AE	0.9675	2.0846	6.0339	0.9924	-1.9994	0.4945
	AB	-0.0325	0.0846	0.0339	-0.0076	0.0006	-0.0055
	AMSE	0.0011	0.0072	0.0011	0.0001	0.0000	0.0000
n=100	AE	0.9753	2.0441	6.0168	0.8860	-1.6365	0.5141
	AB	-0.0247	0.0441	0.0168	-0.1140	0.3635	0.0141
	AMSE	0.0006	0.0019	0.0003	0.0744	1.9252	0.2931
n=300	AE	0.9964	2.0175	6.0067	1.0361	-1.9196	0.5044
	AB	-0.0036	0.0175	0.0067	0.0361	0.0804	0.0044
	AMSE	0.0000	0.0003	0.0000	0.6488	0.0831	0.0040
Sample size	Parameters	$\alpha = 1.5$	$\beta = 2.0$	$\eta = 4.0$	$\gamma_0 = 1.0$	$\gamma_1 = 0.5$	$\gamma_2 = 0.5$
		$\hat{\alpha}$	$\hat{\beta}$	$\hat{\eta}$	$\widehat{\gamma_0}$	$\widehat{\gamma_1}$	$\widehat{\gamma_2}$
n=20	AE	0.9029	1.5146	3.8280	0.1816	0.1653	0.1608
	AB	-0.5971	-0.4854	-0.1720	-0.8184	-0.3347	-0.3392
	AMSE	0.3566	0.2356	0.0296	0.6752	0.1263	0.1205
n=50	AE	0.9403	2.1901	3.6587	0.6245	0.3420	0.3426
	AB	-0.5597	0.1901	-0.3413	-0.3755	-0.1580	-0.1574
	AMSE	0.3133	0.0361	0.1165	0.1410	0.0250	0.0248
n=100	AE	1.2672	2.0408	3.9365	0.8048	0.4201	0.4203
	AB	-0.2328	0.0408	-0.0635	-0.1952	-0.0799	-0.0797

continue to next page

continue from previous page

Sample size	Parameters	$\alpha = 1.5$	$\beta = 2.0$	$\eta = 6.0$	$\gamma_0 = 1.0$	$\gamma_1 = -2.0$	$\gamma_2 = 0.5$
	AMSE	0.0542	0.0017	0.0040	0.0381	0.0064	0.0064
n=300	AE	1.4055	2.0074	3.9887	0.9812	0.4447	0.4732
	AB	-0.0945	0.0074	-0.0113	-0.0188	-0.0553	-0.0268
	AMSE	0.0089	0.0001	0.0001	0.0173	0.0218	0.0019
Sample size	Parameters	$\alpha = 1.5$	$\beta = 2.0$	$\eta = 6.0$	$\gamma_0 = 1.0$	$\gamma_1 = -2.0$	$\gamma_2 = 0.5$
		$\hat{\alpha}$	$\hat{\beta}$	$\hat{\eta}$	$\widehat{\gamma_0}$	$\widehat{\gamma_1}$	$\widehat{\gamma_2}$
n=20	AE	0.8324	2.2891	5.8876	0.8511	-1.7643	0.3723
	AB	-0.6676	0.2891	-0.1124	-0.1489	0.2357	-0.1277
	AMSE	0.4457	0.0836	0.0126	0.0239	0.0565	0.0169
n=50	AE	1.4099	2.0399	5.9864	0.9291	-1.9242	0.4519
	AB	-0.0901	0.0399	-0.0136	-0.0709	0.0758	-0.0481
	AMSE	0.0081	0.0016	0.0002	0.0063	0.0068	0.0026
n=100	AE	1.4042	2.0285	5.9904	0.9566	-1.9594	0.4641
	AB	-0.0958	0.0285	-0.0096	-0.0434	0.0406	-0.0359
	AMSE	0.0092	0.0008	0.0001	0.0034	0.0027	0.0014
n=300	AE	1.4858	2.0033	5.9986	1.0009	-1.9707	0.4946
	AB	-0.0142	0.0033	-0.0014	0.0009	0.0293	-0.0054
	AMSE	0.0002	0.0000	0.0000	0.0009	0.0014	0.0001

TABLE 3. Descriptive statistics for the survival data

Statistics	Mean	Variance	Minimum	Maximum	First quartile	Third quartile
Values	2.0130	4.1661	0.01	8.35	1.1542	1.473

TABLE 4. Fitted distributions and parameter estimations for the survival data

Model	Parameter	MLE	Sd	Lower limit	Upper limit	p-value
Log-normal	$\hat{\lambda}$	0.0650	0.0082	-0.0843	0.2142	0.4742
	$\hat{\alpha}$	1.3407	0.0041	1.2351	1.4463	0.1823
Weibull	$\hat{\lambda}$	0.9470	0.0033	-0.6979	1.5918	0.3437
	$\hat{\alpha}$	1.9640	0.0031	0.3192	2.6089	0.0495
Gompertz	$\hat{\lambda}$	0.1580	0.0091	0.1564	0.1597	0.0000
	$\hat{\alpha}$	1.9993	0.0028	1.9289	2.0886	0.7951
GL	$\hat{\lambda}$	0.6633	0.0034	0.5920	0.8075	0.0006
	$\hat{\alpha}$	1.9997	0.0031	1.8612	2.1733	0.0000

continue to next page

continue from previous page

GTPL	$\hat{\lambda}$	0.4459	0.0036	0.3475	0.5442	0.0160
	$\hat{\alpha}$	0.8078	0.0020	0.2558	1.3598	0.0946
	$\hat{\beta}$	0.0242	0.1290	-0.5667	0.6152	0.0000
GTHPL	$\hat{\lambda}$	0.4145	0.0035	0.3284	0.50068	0.0000
	$\hat{\alpha}$	0.8337	0.0003	0.8003	0.8809	0.0000
	$\hat{\beta}$	3.2823	0.0068	2.5388	4.0258	0.0375
	$\hat{\eta}$	0.3870	0.0056	0.2487	0.4898	0.0122

TABLE 5. Fitted distributions and tests for the survival data

Model	AIC	BIC	CAIC	HQIC
Log-normal	761.7162	754.9656	752.9656	758.9889
Weibull	729.1806	722.4301	720.4301	726.4534
Gompertz	753.5534	746.8029	744.8029	750.8262
GL	924.1786	917.4280	915.4280	921.4513
GTPL	726.6152	716.4893	713.4893	722.5243
GTHPL	724.5317	711.0306	707.0306	719.0772

TABLE 6. Fitted regression models and parameter estimations for the survival data

Model	Parameter	MLE	Sd	Lower limit	Upper limit	P value
GL regression	$\hat{\alpha}$	0.6977	0.2801	0.237	1.1584	0.0127
	$\hat{\gamma}_0$	1.4782	0.1408	1.2467	1.7097	0.0000
	$\hat{\gamma}_1$	0.5309	0.0701	0.4156	0.6461	0.0000
	$\hat{\gamma}_2$	1.4721	0.1392	1.2432	1.701	0.0000
GTPL regression	$\hat{\alpha}$	0.5754	0.1228	0.3734	0.7775	0.0000
	$\hat{\beta}$	0.8949	0.2468	0.4888	1.3009	0.0003
	$\hat{\gamma}_0$	1.4916	0.17	1.212	1.7713	0.0000
	$\hat{\gamma}_1$	0.4272	0.1073	0.2507	0.6036	0.0001
	$\hat{\gamma}_2$	1.3936	0.1965	1.0704	1.7168	0.0000

continue to next page

continue from previous page

GTHPL regression	$\hat{\alpha}$	0.5082	0.1792	0.2135	0.803	0.0046
	$\hat{\beta}$	1.0659	0.5348	0.137	1.9949	0.0462
	$\hat{\eta}$	0.6081	0.2921	0.1277	1.0885	0.0373
	$\hat{\gamma_0}$	1.4985	0.2098	1.1535	1.8435	0.0000
	$\hat{\gamma_1}$	0.4948	0.0763	0.3692	0.6204	0.0000
	$\hat{\gamma_2}$	1.4996	0.1072	1.3233	1.676	0.0000

TABLE 7. Fitted regression models and tests for the survival data

Model	AIC	BIC	CAIC	HQIC
GL regression	714.4258	707.1761	708.8961	712.1846
GTPL regression	713.4114	704.3493	706.4993	710.6098
GTHPL regression	713.3666	702.4921	705.0721	710.0047

CONCLUSIONS

The Gompertz–three–parameter–Lindley distribution, abbreviated as GTHPL, which includes the Gompertz–Lindley distribution and Gompertz–two–parameter–Lindley distribution as special cases, was introduced. This new model is constructed by mixing the Gompertz distribution and the three–parameter Lindley distribution. Several statistical properties have been studied, such as the failure rate function, mean residual life function, r-th moment, skewness and kurtosis. Moreover, the GTHPL regression model based on the proposed distribution has also been developed, offering a useful tool for analyzing survival data with covariates. The maximum likelihood method was used to estimate the parameters of the GTHPL distribution and the GTHPL regression model, and simulation studies were carried out to demonstrate the consistency of the MLEs. Finally, real data analysis further validated the better performance of the GTHPL model compared with established models such as the log–normal, Weibull, Gompertz, Gompertz–Lindley, and Gompertz–two–parameter–Lindley distributions. Additionally, the GTHPL regression model has a better fit than its counterparts do, highlighting its potential as a valuable tool for survival analysis with covariates.

ACKNOWLEDGEMENTS

This work was supported by the Guangdong Basic and Applied Basic Research Foundation (2023A1515110661) and Science and Technology Projects in Guangzhou (2024A04J3550).

REFERENCES

Alizadeh, M., Benkhelifa, L., Rasekhi, M. & Hosseini, B. 2020. The odd log-logistic generalized gompertz distribution: Properties, applications and different methods of estimation. *Communications in Mathematics and Statistics* 8: 295-317.

Al-Omari, A.I., Ciavolino, E. & Al-Nasser, A.D. 2020. Economic design of acceptance sampling plans for truncated life tests using three-parameter Lindley distribution. *Journal of Modern Applied Statistical Methods* 18(2): 2-15.

Benkhelifa, L. 2017. The beta generalized Gompertz distribution. *Applied Mathematical Modelling* 52: 341-357.

Boshi, M.A.A., Abid, S.H. & Al-Noor, N.H. 2019. Generalized Gompertz - Generalized Gompertz distribution. *Journal of Physics-Conference Series* 1234: 012112.

Eghwerido, J.T., Ogbo, J.O. & Omotoye, A.E. 2021. The Marshall-Olkin-Gompertz distribution: Properties and applications. *Statistica* 81: 183-215.

El-Damcese, M.A., Mustafa, A., El-Desouky, B.S. & Mustafa, M.E. 2015. The odd generalized exponential gompertz distribution. *Applied Mathematics* 6: 2340-2353.

Gavrilov, L. & Gavrilova, N. 2001. The reliability theory of aging and longevity. *Journal of Theoretical Biology* 213: 527-545.

Ghitany, M., Alqallaf, F. & Balakrishnan, N. 2014. On the likelihood estimation of the parameters of Gompertz distribution based on complete and progressively type-II censored samples. *Journal of Statistical Computation and Simulation* 84: 1803-1812.

Ghitany, M.E., Atieh, B. & Nadarajah, S. 2008. Lindley distribution and its application. *Mathematics and Computers in Simulation* 78(4): 493-506.

Ghitany, M.E., Aboukhamseen, S.M., Baqer, A.A. & Gupta, R.C. 2019. Gompertz-Lindley distribution and associated inference. *Communications in Statistics-Simulation Computation* 51(5): 2599-2618.

Jafari, A., Tahmasebi, S. & Alizadeh, M. 2014. The beta-gompertz distribution. *Revista Colombiana De Estadistica* 37: 141-158.

Jayakumar, K. & Shabeer, A.M. 2022. On a generalization of Gompertz distribution and its applications. *Journal of the Indian Society for Probability and Statistics* 23: 241-265.

Lenart, A. 2014. The moments of the Gompertz distribution and maximum likelihood estimation of its parameters. *Scandinavian Actuarial Journal* 3: 255-277.

Lenart, A. & Missov, T. 2016. Goodness-of-fit tests for the Gompertz distribution. *Communications in Statistics - Theory and Methods* 45: 2920-2937.

Lindley, D.V. 1958. Fiducial distributions and Bayes theorem. *Journal of the Royal Statistical Society Series B* 20: 102-107.

Marshall, A.W. & Olkin, I. 1997. A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families. *Biometrika* 84: 641-652.

Ou, X.H., Lu, H.Z. & Kong, J.S. 2022. The structural properties of the Gompertz-Two-Parameter-Lindley distribution and associated inference. *Open Mathematics* 20(1): 1581-1593.

Shen, L.J., Zeng, Q., Guo, P., Huang, J.J., Li, C.F., Pan, T., Chang, B.Y., Wu, N., Yang, L.W., Chen, Q.F., Huang, T., Li, W. & Wu, P.H. 2018. Dynamically prognosticating patients with hepatocellular carcinoma through survival paths mapping based on time-series data, *Nature Communications* 9: 2230.

Shaked, M. & Shanthikumar, J.G. 2007. *Stochastic Orders*. New York: Springer.

Shama, M.S., Dey, S., Altun, E. & Afify, A.Z. 2022. The Gamma-Gompertz distribution: Theory and applications. *Mathematics and Computers in Simulation* 193: 689-712.

Shanker, R. & Sharma, S. 2013. A two-parameter Lindley distribution for modeling waiting and survival times data. *Applied Mathematics* 4: 363-368.

Shanker, R., Shukla, K.K. & Mishra, A. 2017. A three-parameter weighted Lindley distribution and its applications to model survival time. *Statistics in Transition New Series* 18(2): 291-310.

Shaker, R., Shukla, K.K., Shanker, R. & Leonida, T.A. 2017. A three-parameter Lindley distribution. *American Journal of Mathematics and Statistics* 7(1): 15-26.

Thamer, M.K. & Zine, R. 2023. Statistical properties and estimation of the three-parameter Lindley distribution with application to COVID-19 data. *Sains Malaysiana* 52(2): 669-682.

Xi, Y.H., Lu, H.Z. & Liang, F. 2024. On a new transmuted three-parameter Lindley distribution and its applications. *Sains Malaysiana* 53(6): 1427-1440.

*Corresponding author; email: luhz@gzhu.edu.cn