International Journal of Statistical Distributions and Applications
2024, Vol. 10, No. 3, pp. 48-59
https://doi.org/10.11648/j.ijsd.20241003.11

Research Article

Transmuted Power Gumbel Distribution: Estimation and Applications

Ahmed Ali Hurairah[1, *], Nasr Tawfiq Almazaqi[2]

[1]Department of Statistics, Sana'a University, Sana'a, Yemen
[2]Department of Mathematics, Sana'a University, Sana'a, Yemen

Abstract

In recent years, generalized distributions have been widely studied in statistics as they possess flexibility in applications. This is justified because the traditional distributions often do not provide good fit in relation to the real data set studied. This paper develops a Power Gumbel distribution using the quadratic rank transmutation map (QRTM). The new generalization is called the transmuted Power-Gumbel distribution. Various mathematical properties of this distribution including moments, moment generating function, quantile function, mean deviation and order statistics were also studied. These features support the legitimacy and robustness of the proposed distribution. The maximum likelihood method is used for estimating the model parameters, and the finite sample performance of the estimators are assessed by simulation studies indicating that their precision improves with larger sample sizes. The asymptotic confidence intervals for the parameters are also obtained based on asymptotic variance-covariance matrix. Finally, the usefulness of the proposed model is illustrated in an application to two real data sets and conclude that the four-parameter transmuted Power Gumbel distribution provides better fit than the other five models.

Keywords:

Gumbel Distribution, Transmuted Power Gumbel Distribution, Parameter Estimation, Asymptotic Confidence Intervals, Quadratic Rank Transmutation

1. Introduction

The Gumbel distribution is perhaps the most widely applied statistical distribution for problems in engineering. It is also known as the extreme value distribution of type I. Recent developments focus on new techniques for building meaningful distributions. These include the two-piece approach introduced by [17], Azzalini and Capitanio [5] studied the distributions generated by perturbation of symmetry with emphasis on a multivariate skew t distribution, and the generator approach pioneered by [14]. Many researchers have worked using generalizations technique, a few to mention are [3, 6, 7-11, 16, 21-23, 25, 26].

The aim of this paper is to introduce a new generalization to the Gumbel distribution using the transmutation map approach introduced by [24]. The new model which generalizes the Power Gumbel (PG) distribution is referred to as the transmuted Power Gumbel (TPG) distribution. The Gumbel (G) distribution is a very popular statistical distribution due to its extensive applicability in several areas and its wide applications has been reported by [19]. The applicability of GD in the field of flood frequency analysis, network, space,

*Corresponding author: Ahmed.Hurairah@su.edu.ye (Ahmed Ali Hurairah)

Received: 6 August 2024; **Accepted:** 5 September 2024; **Published:** 23 September 2024

software reliability, structural and wind engineering are reported by [8]. Exponentiated Gumbel (EG) distribution, introduced by [20] based on Gumbel (G) distribution and illustrated its applicability in the area of global warming modeling, rainfall modeling, wind speed modeling etc. Due to its wide applicability in different fields of science, the generalization of Gumbel Distribution has become important. The cumulative distribution function (cdf) of the Power Gumbel distribution is given by

$$F(x) = \int_0^{G(x)} f(y)dy = \left(\mathbf{e}^{-e^{-\frac{x-\mu}{\sigma}}}\right)^{\frac{1}{\alpha}} \qquad (1)$$

where σ is a scale parameter, α is a shape parameter and μ is a location parameter, the corresponding probability density function (pdf) is given by

$$f(x) = \frac{1}{\alpha\,\sigma}\, e^{-\frac{x-\mu}{\sigma}} \left(e^{-e^{-\frac{x-\mu}{\sigma}}}\right)^{\frac{1}{\alpha}} \qquad (2)$$

According to the Quadratic Rank Transmutation Map, (QRTM), approach a random variable X is said to have transmuted distribution if its cumulative distribution function cdf is given by

$$F(x) = (1+\lambda)G(x) - \lambda[G(x)]^2 \qquad (3)$$

where $G(x)$ is the cdf of the base distribution, which on differentiation yields

$$f(x) = g(x)[(1+\lambda) - 2\lambda G(x)], \quad |\lambda| \le 1 \qquad (4)$$

where $f(x)$ and $g(x)$ are the corresponding pdfs associated with cdfs $F(x)$ and $G(x)$ respectively. Note that when $\lambda=0$, the base distribution will be obtained. Based on the above generalization, many authors [1, 4, 12, 13, 18] have dealt with this generalization of some known distributions. Afify et al. [2] derived the Transmuted Complementary Weibull Geometric distribution. The rest of the paper is organized as follows. In Section 2, we introduce the new distribution. In Section 3, we obtain some mathematical properties of the TPG distribution including, moments and moment generating functions, quantile, mean deviations, Renyi entropies and order statistics. In Section 4, we discuss the estimation problem using the maximum likelihood estimation method and obtain the observed information matrix. The asymptotic confidence intervals for the parameters are also obtained based on asymptotic variance-covariance matrix. The finite sample performance of the estimators and approximate $100(1-\vartheta)$ % confidence intervals are assessed by simulation in Section 5. Two illustrative applications based on real data sets are investigated in section 6. Finally, concluding remarks are presented in section 7.

2. The Transmuted Power Gumbel Distribution

In this section, we studied the transmuted Power-Gumbel (TPG) distribution. Now by inserting (1) into (3), we obtain the cdf of transmuted Power-Gumbel (TPG) distribution

$$F(x) = \left(e^{-e^{-\frac{x-\mu}{\sigma}}}\right)^{\frac{1}{\alpha}} \left(1+\lambda - \lambda\left(e^{-e^{-\frac{x-\mu}{\sigma}}}\right)^{\frac{1}{\alpha}}\right). \qquad (5)$$

The corresponding probability density function pdf is given by

$$f(x) = \frac{1}{\alpha\,\sigma}\, e^{-\frac{x-\mu}{\sigma}} \left(e^{-e^{-\frac{x-\mu}{\sigma}}}\right)^{\frac{1}{\alpha}} \left[1+\lambda - 2\lambda\left(e^{-e^{-\frac{x-\mu}{\sigma}}}\right)^{\frac{1}{\alpha}}\right] \qquad (6)$$

where μ is a location parameter, σ is a scale parameter, α is shape parameter and λ is a transmuted parameter.

Figures 1 and 2 illustrate the graphical behavior of the pdf and cdf of transmuted Power-Gumbel distribution for selected values of the parameters.

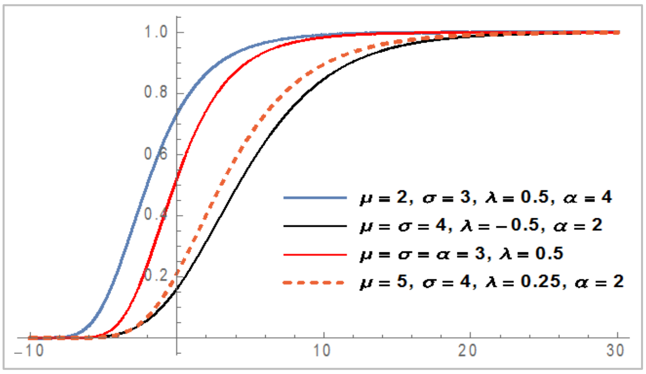

Figure 1. *The CDF of the TPG distribution.*

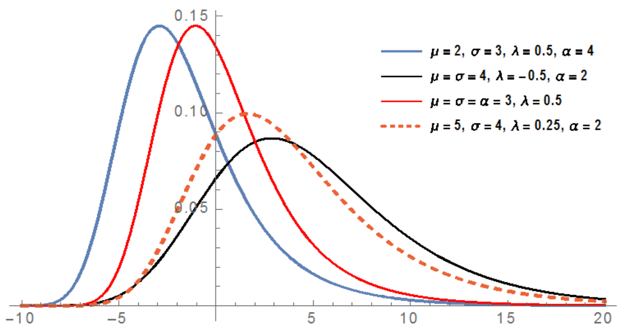

Figure 2. *The PDF of theTPG distribution.*

3. Mathematical Properties

In this section we provide some mathematical properties of the TPG distribution including the moments, moment generating function, quantiles, mean deviations, Rényi entropies and order statistic.

3.1. Moments and Moment Generating Function

It is imperative to derive the moments when a new distribution is proposed. They play a significant role in statistical analysis, particularly in applications. If X has Transmuted Power Gumbel distribution (μ, σ, α, λ) then its rth moment of TPG can be written as

$$E(X^r) = \int_{-\infty}^{\infty} X^r \frac{1}{\alpha\,\sigma}\ e^{-\frac{x-\mu}{\sigma}} (e^{-e^{-\frac{x-\mu}{\sigma}}})^{\frac{1}{\alpha}} \left[1 + \lambda - 2\lambda(e^{-e^{-\frac{x-\mu}{\sigma}}})^{\frac{1}{\alpha}}\right] \tag{7}$$

$$E(X^r) = 2^{-r} e^{-\frac{2\mu}{\alpha\sigma\log}} (-\alpha\sigma\log)^r \left((\lambda+1)2^r e^{\frac{\mu}{\alpha\sigma\log}} \Gamma\left(r+1, -\frac{\mu}{\alpha\log\sigma}\right) - \lambda\Gamma\left(r+1, -\frac{\mu}{\alpha\log\sigma}\right)\right) \tag{8}$$

In particular,

$$E(X) = \mu - \frac{1}{2}\alpha\log(2+\lambda)\sigma$$

$$E(X^2) = \mu^2 - \alpha\log(2+\lambda)\mu\sigma + \frac{1}{2}\alpha^2\log^2(4+3\lambda)\sigma^2$$

$$E(X^3) = \frac{1}{4}(4\mu^3 - 6\alpha\log(2+\lambda)\mu^2\sigma + 6\alpha^2\log^2(4+3\lambda)\mu\sigma^2 - 3\alpha^3\log^3(8+7\lambda)\sigma^3)$$

$$E(X^4) = \mu^4 - 2\alpha\log(2+\lambda)\mu^3\sigma + 3\alpha^2\log^2(4+3\lambda)\mu^2\sigma^2 - 3\alpha^3\log^3(8+7\lambda)\mu\sigma^3 + \frac{3}{2}\alpha^4\log^4(16+15\lambda)\sigma^4$$

The variance, skewness, and kurtosis measures can now be calculated using the relations

$$Var(X) = E(X^2) - [E(X)]^2$$

$$\text{Skewness}(X) = \frac{E(X^3) - 3E(X)E(X^2) + 2[E(X)]^3}{[Var(X)]^{\frac{3}{2}}}$$

$$\text{Kurtosis}(X) = \frac{E(X^4) - 4E(X)E(X^3) + 6E(X^2)[E(X)]^2 - 3[E(X)]^4}{[Var(X)]^2}$$

Similarly, the moment generating function of X obtained as below:

$$M_x(t) = E(\mathbf{e}^{tx}) = \int_{-\infty}^{\infty} e^{t\,x} f(x)\,dx$$

$$M_x(t) = e^{t\mu} 2^{\sigma(-t)} a^{\sigma t}((\lambda+1)2^{\sigma t} - \lambda)\Gamma(t\sigma+1)$$

3.2. Mean and Median Deviations

The amount of scatter in a population is evidently measured to some extent by the totality of deviations from the mean and median. If X has a TPG distribution, then we can derive the mean deviation about the mean $\mu = E(X)$ and about the median M as:

$$\delta_1(x) = \int_{-\infty}^{\infty} |x - \mu| f(x)dx$$

$$\delta_2(x) = \int_{-\infty}^{\infty} |x - \mathrm{M}| f(x)dx.$$

The mean of the distribution is obtained from (8), and the median is obtained by solving the equation:

$$\left(e^{-e^{-\frac{x-\mu}{\sigma}}}\right)^{\frac{1}{\alpha}} \left(1 + \lambda - \lambda\left(e^{-e^{-\frac{x-\mu}{\sigma}}}\right)^{\frac{1}{\alpha}}\right) = \frac{1}{2} \tag{9}$$

These measures can be calculated using the relationships that:

$$\delta_1(x) = \int_{-\infty}^{\mu} |\mu - x| f(x)dx + \int_{\mu}^{\infty} |x - \mu| f(x)dx$$

$$= 2\int_0^{\mu} |\mu - x| f(x)dx = 2\{\mu F(\mu) - \int_0^{\mu} x f(x)dx\}$$

$$\delta_1(x) = 2\{\mu F(\mu) - J(\mu)\}$$

$$\delta_2(x) = \mu - 2J(\mu)$$

where $J(t) = \int_{-\infty}^{t} x f(x)dx$. From (6) we have:

$$J(t) = \int_{-\infty}^{t} x f(x)dx$$

$$= -\left(e^{-e^{\frac{-t+\mu}{\sigma}}}\right)^{\frac{1}{\alpha}}\left(t\left(-1+\left(-1+\left(e^{-e^{\frac{-t+\mu}{\sigma}}}\right)^{\frac{1}{\alpha}}\right)\lambda\right)+e^{\frac{2e^{\frac{-t+\mu}{\sigma}}}{\alpha}}\left(e^{-e^{\frac{-t+\mu}{\sigma}}}\right)^{\frac{1}{\alpha}}\lambda\sigma \mathrm{Ei}\left[-\frac{2e^{\frac{-t+\mu}{\sigma}}}{\alpha}\right]-e^{\frac{e^{\frac{-t+\mu}{\sigma}}}{\alpha}}(1+\lambda)\sigma\mathrm{Ei}\left[-\frac{e^{\frac{-t+\mu}{\sigma}}}{\alpha}\right]\right), \quad (10)$$

where Ei is the exponential integral function Ei(z) defined by $\mathrm{Ei}(z) = -\int_{-z}^{\infty} \frac{e^{-t}}{t} dt$

3.3. Quantiles and Random Number Generator

The quantile function plays a key role in simulating random samples from a given distribution. The characteristics of a distribution such as the median, kurtosis and skewness can also be described using the quantile function. The quantiles function of the TPG distribution is the real solution of $F(x_q) = q$ for $0 \le q \le 1$. The quantiles of the Transmuted power Gumbel distribution are obtained from cdf (5) as:

$$x_p = \mu - \sigma Log\left[Log\left[2^{\alpha}\left(\frac{1+\lambda-\sqrt{1+2\lambda-4U\lambda+\lambda^2}}{\lambda}\right)^{-\alpha}\right]\right] \quad (11)$$

where U is a uniform distribution with $(0,1)$.

3.4. Rényi Entropies

An entropy of a random variable x is a measure of variation of the uncertainty. Rényi entropy is defined by:

$$\mathcal{J}_R(\gamma) = \frac{1}{1-\gamma} log[\int f^{\gamma}(x)dx]$$

where $\gamma > 0$ and $\gamma \neq 1$. For the TPG distribution pdf given by (6)

$$\mathcal{J}_R(\gamma) = \frac{1}{a^{\gamma}\sigma^{\gamma}}\int_{-\infty}^{\infty}\left(e^{-\frac{x-\mu}{\sigma}}\right)^{\gamma}\left(e^{-e^{-\frac{x-\mu}{\sigma}}}\right)^{\frac{1}{\alpha}}\left[1+\lambda-2\lambda\left(e^{-e^{-\frac{x-\mu}{\sigma}}}\right)^{\frac{1}{\alpha}}\right]^{\gamma} dx$$

On substituting $y = e^{-\frac{x-\mu}{\sigma}}$, the right-hand side reduces to

$$= \int_0^{\infty}\frac{\sigma^{\gamma}}{y^{\gamma}}(y)^{\gamma}(e^{-y})^{\frac{1}{\alpha}}\left[1+\lambda-2\lambda(e^{-y})^{\frac{1}{\alpha}}\right]^{\gamma} dy = \frac{\alpha\,\Gamma(\gamma)\left(\frac{1+\lambda}{\alpha}\right)^{\gamma} 2F1[-\gamma,\gamma,1+\gamma,\frac{2\lambda}{1+\lambda}]}{\gamma}$$

The Rényi entropy, takes the expression

$$\mathcal{J}_R(\gamma) = \frac{\log\left(\alpha\Gamma(\gamma)\left(\frac{1+\lambda}{\alpha}\right)^{\gamma} 2F1\left[-\gamma,\gamma,1+\gamma,\frac{2\lambda}{1+\lambda}\right]\right)}{\gamma-1}, \tag{12}$$

where, $\widetilde{2F_1}[a,b,c,z]$ is the hypergeometric function calculated as.

$$\widetilde{2F_1} = \sum_{k=0}^{\infty} \frac{a_k b_k}{k!\, c_k} z^k.$$

3.5. Order Statistics

Order statistics make their appearance in many areas of statistical theory and practice. Suppose $X_1, X_2, X_3, \dots \dots, X_n$ is a random sample from the TPG distribution. Let $X_{i:n}$ denote the ith order statistics. The pdf of $X_{i:n}$ can be expressed as

$$f_{i:n}(x) = \frac{n!}{(i-1)!(n-i)!} (F(x))^{i-1} (1-F(x))^{n-i} f(x)$$

Let $X_1, X_2, X_3, \dots \dots, X_n$ be independently identically distributed order random variables from the (TPG) distribution having first and last order probability density function are given by the following:

$$f_{1:n}(x) = n\ (1-F(x))^{n-1} f(x)$$

$$= n\ \left(1-(1+\lambda)\,(e^{-e^{-\frac{x-\mu}{\sigma}}})^{\frac{1}{\alpha}} - \lambda(e^{-e^{-\frac{x-\mu}{\sigma}}})^{\frac{2}{\alpha}}\right)^{n-1} \left[\frac{1+\lambda}{\alpha\,\sigma}\ e^{-\frac{x-\mu}{\sigma}} \left(e^{-e^{-\frac{x-\mu}{\sigma}}}\right)^{\frac{1}{\alpha}} - \frac{2\lambda}{\alpha\,\sigma}\ e^{-\frac{x-\mu}{\sigma}} \left(e^{-e^{-\frac{x-\mu}{\sigma}}}\right)^{\frac{2}{\alpha}}\right] \tag{13}$$

$$f_{n:n}(x) = n\ (F(x))^{n-1} f(x)$$

$$= n\ \left((1+\lambda)\left(e^{-e^{-\frac{x-\mu}{\sigma}}}\right)^{\frac{1}{\alpha}} - \lambda\left(e^{-e^{-\frac{x-\mu}{\sigma}}}\right)^{\frac{2}{\alpha}}\right)^{n-1} \left[\frac{1+\lambda}{\alpha\,\sigma}\ e^{-\frac{x-\mu}{\sigma}} \left(e^{-e^{-\frac{x-\mu}{\sigma}}}\right)^{\frac{1}{\alpha}} - \frac{2\lambda}{\alpha\,\sigma}\ e^{-\frac{x-\mu}{\sigma}} \left(e^{-e^{-\frac{x-\mu}{\sigma}}}\right)^{\frac{2}{\alpha}}\right] \tag{14}$$

3.6. Survivor Function

The survivor function S(x) is given by

$$S(x) = p(X \geq x) = 1 - p(X \leq x) = 1 - F(x)$$

Substituting the value of F(x) from (5), we get

$$S(x) = 1 - \left(e^{-e^{-\frac{x-\mu}{\sigma}}}\right)^{\frac{1}{\alpha}} \left(1 + \lambda - \lambda\left(e^{-e^{-\frac{x-\mu}{\sigma}}}\right)^{\frac{1}{\alpha}}\right)$$

Figure 3 is the plot of the survival function of the TPG for various sets of parameter values.

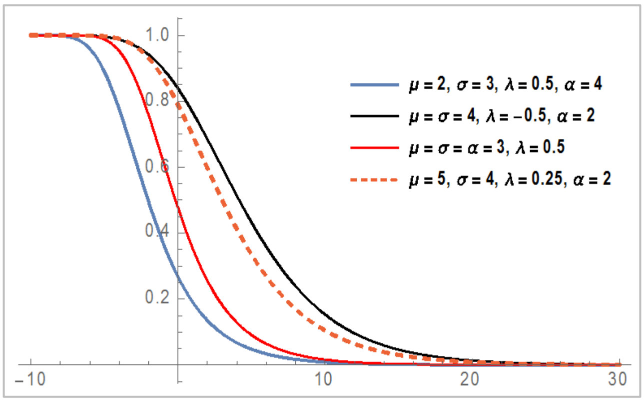

Figure 3. *Survival function of the TPG plotted for various parameter values.*

3.7. Hazard Rate Function

The hazard rate function h(x) is given by $h(x)=\frac{f(x)}{S(x)}$. Substituting the values of f(x) and F(x) from (5) and (6) respectively, and simplifying we get

$$h(x)=\frac{\frac{1+\lambda}{\alpha\,\sigma}\,e^{-\frac{x-\mu}{\sigma}}\left(e^{-e^{-\frac{x-\mu}{\sigma}}}\right)^{\frac{1}{a}}-\frac{2\lambda}{\alpha\,\sigma}\,e^{-\frac{x-\mu}{\sigma}}\left(e^{-e^{-\frac{x-\mu}{\sigma}}}\right)^{\frac{2}{\alpha}}}{1-\left(e^{-e^{-\frac{x-\mu}{\sigma}}}\right)^{\frac{1}{\alpha}}\left(1+\lambda-\lambda\left(e^{-e^{-\frac{x-\mu}{\sigma}}}\right)^{\frac{1}{\alpha}}\right)}$$

4. Parameter Estimation

The maximum likelihood estimators (MLEs) enjoy desirable properties and can be used for constructing confidence intervals for the model parameters. So, we consider the estimation of the unknown parameters for this distribution by maximum likelihood. Let $X_1, X_2 ..., X_n$ be a random sample from the TPG distribution with observed values $x_1, x_2 ..., x_n$ and $\Theta=(\mu,\sigma,\alpha,\lambda)^T$ be parameter vector. The likelihood function for Θ may be expressed as

$$L(\Theta)=\prod_{i=1}^{n}\frac{1}{\alpha\,\sigma}\;e^{-\frac{x_i-\mu}{\sigma}}\left(e^{-e^{-\frac{x_i-\mu}{\sigma}}}\right)^{\frac{1}{\alpha}}\left[1+\lambda-2\lambda\left(e^{-e^{-\frac{x_i-\mu}{\sigma}}}\right)^{\frac{1}{\alpha}}\right]$$

Therefore, the log-likelihood function for Θ becomes

$$l(\Theta)=-nLn\alpha-nLn\sigma-\sum\frac{x-\mu}{\sigma}-\frac{1}{\alpha}\sum e^{-\frac{x-\mu}{\sigma}}+Ln\prod_{i=1}^{n}\left[1+\lambda-2\lambda\left(e^{-e^{-\frac{x-\mu}{\sigma}}}\right)^{\frac{1}{\alpha}}\right] \tag{15}$$

The MLEs of $\mu,\sigma,\alpha,\lambda$ say $\hat{\mu},\hat{\sigma},\hat{\alpha}$ *and* $\hat{\lambda}$, respectively, can be worked out by the solutions of the system of equations obtained by letting the first partial derivatives of the total log-likelihood equal to zero with respect to $\hat{\mu},\hat{\sigma},\hat{\alpha}$ *and* $\hat{\lambda}$. Therefore, the system of equations is as follows:

$$\frac{\partial\ln l(\Theta)}{\partial\mu}=\sum\frac{1}{\sigma}-\frac{1}{\alpha}\sum e^{-\frac{x-\mu}{\sigma}}\left(\frac{1}{\sigma}\right)+\sum\frac{2\lambda\left(e^{-e^{-\frac{x-\mu}{\sigma}}}\right)^{\frac{1}{\alpha}}\left(e^{-\frac{x-\mu}{\sigma}}\right)\left(\frac{1}{\sigma}\right)}{1+\lambda-2\lambda\left(e^{-e^{-\frac{x-\mu}{\sigma}}}\right)^{\frac{1}{\alpha}}} \tag{16}$$

$$\frac{\partial\ln l(\Theta)}{\partial\sigma}=-\frac{n}{\sigma}+\sum\frac{x-\mu}{\sigma^2}-\frac{1}{\alpha}\sum e^{-\frac{x-\mu}{\sigma}}\left(\frac{x-\mu}{\sigma^2}\right)+\sum\frac{2\lambda\left(e^{-e^{-\frac{x-\mu}{\sigma}}}\right)^{\frac{1}{\alpha}}\left(e^{-\frac{x-\mu}{\sigma}}\right)\left(\frac{x-\mu}{\sigma^2}\right)}{1+\lambda-2\lambda\left(e^{-e^{-\frac{x-\mu}{\sigma}}}\right)^{\frac{1}{\alpha}}} \tag{17}$$

$$\frac{\partial\ln l(\Theta)}{\partial\alpha}=\frac{1}{\alpha^2}\sum e^{-\frac{x-\mu}{\sigma}}-\frac{n}{\alpha}-\sum\frac{2\lambda\left(e^{-e^{-\frac{x-\mu}{\sigma}}}\right)^{\frac{1}{\alpha}}\left(\frac{1}{\alpha^3}\right)\left(e^{-e^{-\frac{x_i-\mu}{\sigma}}}\right)}{1+\lambda-2\lambda\left(e^{-e^{-\frac{x-\mu}{\sigma}}}\right)^{\frac{1}{\alpha}}} \tag{18}$$

$$\frac{\partial\ln l(\Theta)}{\partial\lambda}=\sum\frac{1-2\left(e^{-e^{-\frac{x-\mu}{\sigma}}}\right)^{\frac{1}{\alpha}}}{1+\lambda-2\lambda\left(e^{-e^{-\frac{x-\mu}{\sigma}}}\right)^{\frac{1}{\alpha}}} \tag{19}$$

The solutions of nonlinear equations (16), (17), (18) and (19) are complicated to obtain, therefore an iterative procedure is applied to solve these equations numerically.

Under certain regularity conditions, the centered form of the MLE, $\sqrt{n}(\widehat{\Theta}-\Theta)$, is asymptotically distributed as Normal $(0, I^{-1}(\Theta))$, where $I(\Theta)$ is the information matrix, given by $I(\Theta)=E\left[\frac{\partial^2 l(\Theta)}{\partial\Theta_i\partial\Theta_j}\right]$, and can be approximated by $I(\widehat{\Theta})$. Then, based on $(\widehat{\Theta}-\Theta)\xrightarrow{d} N_4(0, I^{-1}(\Theta))$, one can carry out confidence regions for functions of Θ. Therefore, approximate $100(1-\gamma)\%$

$$\hat{\mu} \mp Z_{\frac{\gamma}{2}}\sqrt{\hat{I}_{\mu\mu}},\ \hat{\sigma} \mp Z_{\frac{\gamma}{2}}\sqrt{\hat{I}_{\sigma\sigma}},\ \hat{\alpha} \mp Z_{\frac{\gamma}{2}}\sqrt{\hat{I}_{\alpha\alpha}},\ \hat{\lambda} \mp Z_{\frac{\gamma}{2}}\sqrt{\hat{I}_{\lambda\lambda}}.$$

where Z_γ is the upper $100\gamma-th$ percentile of the standard normal distribution.

5. Simulation Study

Here, we assess the finite sample behaviors of the MLEs for the four-parameter TPG distributions. The assessment of the finite sample behavior of the MLEs for this distribution was based on the following:

1. Use the inversion method to generate two thousand samples of size n from the TPG distribution, i.e. generate values of:

$$X=\mu-\sigma\mathrm{Log}\left[\mathrm{Log}\left[2^{\alpha}\left(\frac{1+\lambda-\sqrt{1+2\lambda-4U\lambda+\lambda^2}}{\lambda}\right)^{-\alpha}\right]\right]$$

2. Compute the MLEs for the two thousand samples, say $(\hat{\mu}_i, \hat{\sigma}_i, \hat{\alpha}_i, \hat{\lambda}_i)$ for $i=1,2,\ldots,2000$.
3. Compute the bias, mean squared errors, standard errors and 95% confidence limits (L, U and T) for two thousand samples. The asymptotic variance (ASV) and 95% confidence intervals are computed by inverting the observed information matrix. Bias and MSE are given by:

$$Bias(\widehat{\Theta})=\frac{\sum_{i=1}^{2000}(\widehat{\Theta}_i-\Theta)}{2000}$$

$$MSE(\widehat{\Theta})=\frac{\sum_{i=1}^{2000}(\widehat{\Theta}-\Theta)^2}{2000}$$

for $\Theta=(\mu,\sigma,\alpha,\lambda)$.

4. We repeat these steps 2000 times (iteration) for n= 10, 30, 50, 80, 100, 150, and 200, so computing $Bias(\widehat{\Theta})$, $MSE(\widehat{\Theta})$, $ASV(\widehat{\Theta})$ and 95% confidence limits (L, U and T) for $\Theta=\mu,\sigma,\alpha,\lambda$.

The average estimates, along with the bias, mean squared error and asymptotic variance are presented in Table 1. In Table 2, the average 95% confidence intervals are reported.

Table 1. *Average MLEs of the parameters and the corresponding mean squared errors (in parenthesis).*

		Sample size						
	n	**10**	**30**	**50**	**80**	**100**	**150**	**200**
$\hat{\mu}$	Estimate	0.5000360	0.5000249	0.5000234	0.5000236	0.5000238	0.5000241	0.5000225
$\hat{\sigma}$		0.3999766	0.3999840	0.3999850	0.3999848	0.3999846	0.3999843	0.3999855
$\hat{\alpha}$		1.2999519	1.299993	1.2999933	1.2999933	1.2999683	1.2999678	1.2999931
$\hat{\lambda}$		0.2022122	0.2019203	0.2018955	0.2019614	0.2019951	0.2020381	0.2018951
$\hat{\mu}$	Bias	0.3499960	0.1166660	0.0700120	0.0437510	0.0349998	0.0233332	0.0174999
$\hat{\sigma}$		0.2600020	0.0866670	0.0520110	0.0325010	0.0260002	0.0173334	0.0130001
$\hat{\alpha}$		0.1700050	0.0566680	0.0340010	0.0212510	0.0170003	0.0113335	0.0035111
$\hat{\lambda}$		0.0797790	0.0266030	0.0159620	0.0099750	0.0079800	0.0053197	0.0039905
$\hat{\mu}$	MSE	0.1224975	0.0136109	0.0048999	0.0019140	0.0012250	0.0005444	0.0003062
$\hat{\sigma}$		0.0676012	0.0075112	0.0027040	0.0010563	0.0006760	0.0003004	0.0001690
$\hat{\alpha}$		0.0289016	0.0032112	0.0011560	0.0004516	0.0002890	0.0001284	0.0000123
$\hat{\lambda}$		0.0063647	0.0007077	0.0002548	0.0000995	0.0000637	0.0000283	0.0000159
$\hat{\mu}$	ASV	2.8648052	0.9453186	0.5640477	0.3511622	0.2802932	0.1858870	0.1395571

	Sample size						
n	10	30	50	80	100	150	200
$\hat{\sigma}$	0.1689974	0.0563327	0.0337997	0.0211248	0.0168998	0.0112665	0.0084499
$\hat{\alpha}$	0.0089971	0.0029993	0.0017996	0.0011248	0.0008998	0.0005999	0.0000768
$\hat{\lambda}$	0.0001071	0.0000320	0.0000189	0.0000117	9.2992E-6	6.1976E-6	4.6288E-6

From Table 1 it is observed that as the sample size increases, the average biases, asymptotic variance and the mean squared errors decrease. This verifies the consistency properties of the estimates.

Table 2. *Average 95% confidence intervals for the parameters.*

n	$\hat{\mu}$			$\hat{\sigma}$			$\hat{\alpha}$			$\hat{\lambda}$		
	L	U	T	L	U	T	L	U	T	L	U	T
10	(0.073	0.051	0.124)	(0.056	0.046	0.102)	(0.062	0.048	0.110)	(0.056	0.049	0.105)
30	(0.065	0.045	0.110)	(0.048	0.04	0.088)	(0.051	0.041	0.092)	(0.044	0.034	0.078)
50	(0.053	0.038	0.091)	(0.042	0.036	0.078)	(0.036	0.03	0.066)	(0.036	0.028	0.064)
80	(0.042	0.031	0.073)	(0.035	0.031	0.066)	(0.036	0.028	0.064)	(0.030	0.024	0.054)
100	(0.035	0.025	0.060)	(0.031	0.027	0.058)	(0.030	0.024	0.054)	(0.028	0.022	0.050)
150	(0.032	0.022	0.054)	(0.029	0.023	0.052)	(0.028	0.021	0.049)	(0.026	0.022	0.048)
200	(0.031	0.018	0.049)	(0.027	0.021	0.044)	(0.027	0.02	0.047)	(0.025	0.022	0.047)

Table 2 shows that as the sample size increases, the average confidence lengths decrease and the intervals tend towards symmetry.

6. Applications

Here, we provide applications to two real data sets to show how the TPG distribution can be applied in practice. For this aim, the TPG distribution is compared with other competitive distributions. In these applications, the model parameters are estimated by the method of maximum likelihood with their corresponding standard errors of the parameters. The Akaike information criteria (AIC), Consistent Akaikes Information Criterion (CAIC), Bayesian information criterion (BIC) and Hannan-Quinn information criterion (HQIC) statistics are computed to compare the fitted models. In general, the smaller values of these statistics, the better the fit to the data.

6.1. The First Real Data

The first real data set included survival times (in months) for a sample of (131) women with breast cancer residing at Dar Al-Hayat of the National Cancer Control Foundation – Sana’a during the period from February to December 2022. The data are:

Table 3. *The first data.*

8	1	1	4	2	4	2	2	1	5
1	5	1	2	2	2	2	2	6	6
9	1	3	2	5	3	2	6	1	6
3	1	2	1	2	1	11	1	1	4

2	5	3	2	1	7	2	9	1	5
2	1	1	1	2	2	2	2	7	1
1	1	1	11	1	1	3	1	1	3
1	3	3	11	3	5	1	2	1	3
1	11	6	1	1	2	5	1	3	3
1	2	6	1	1	3	4	2	2	2
1	2	8	4	4	9	3	6	2	1
1	5	8	8	1	5	3	4	1	1
1	1	3	2	8	4	4	5	1	1
1									

The descriptive statistics of the data are shown in Table 4 below.

***Table 4**. Descriptive Statistics of the first data.*

N	Minimum	Mean	Median	SD	Variance	Skewness	Kurtosis	Maximum
131	1	3.0916	2	2.54936	6.49924	1.44367	4.44689	11

We compare the four-parameter transmuted Power Gumbel (TPG) distribution, three parameter Power Gumbel (PG) distribution, three parameters transmuted Gumbel (TG) distribution, two parameters transmuted Power (TP) distribution, two parameter Gumbel (G) distribution and one parameter Power (P) distribution fitted to the data. In general, the smaller the values of these statistics, the better the fit to the data.

The estimates of the parameters and their standard errors (SEs) are listed in Tables 5 and 8. The values of the statistics Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Hannan–Quinn Information Criterion (HQIC) and Consistent Akaike Information Criterion (CAIC) are also given in Tables 5 and 8.

***Table 5**. MLEs (standard errors in parentheses) and the measures AIC, BIC, HQIC and CAIC of the first data.*

Distribution	$\hat{\mu}$	$\hat{\sigma}$	$\hat{a}$	$\hat{\lambda}$	AIC	BIC	HQIC	CAIC
TPG	0.5072283 (1.519E-10)	0.3950362 (1.15E-10)	1.2977409 (1.524E-11)	0.243313 (2.168E-11)	542.36787	553.86866	547.04115	542.68533
PG	1.3520571 (0.000037)	1.2980909 (0.0000354)	0.9958667 (0.0000191)	-	667.10618	675.73177	670.61114	667.29515
TG	0.3547992 (0.0424285)	1.6546951 (1.2150712)	-	1.1041474 (0.0107192)	621.60111	630.2267	625.10607	621.79008
TP	-	-	0.3994155 (0.0006942)	0.1948232 (0.1195048)	662.25883	668.00922	664.59547	662.35258
G	0.455232 (16.002212)	1.4050549 (17.3974)	-	-	704.98438	710.73478	707.32102	705.07813
P	2.0006681 (2.6420308)	-	-	-	760.24241	768.868	763.74737	760.43139

6.2. The Second real Data

The second real data set consists of 100 observations on waiting time (in minutes) before the customer received service in a bank which is extracted from [15]. The data are:

Table 6. *The second real data.*

0.8	0.8	1.3	1.5	1.8	1.9	1.9	2.1	2.6	2.7
2.9	3.1	3.2	3.3	3.5	3.6	4.0	4.1	4.2	4.2
4.3	4.3	4.4	4.4	4.6	4.7	4.7	4.8	4.9	4.9
5.0	5.3	5.5	5.7	5.7	6.1	6.2	6.2	6.2	6.3
6.7	6.9	7.1	7.1	7.1	7.1	7.4	7.6	7.7	8.0
8.2	8.6	8.6	8.6	8.8	8.8	8.9	8.9	9.5	9.6
9.7	9.8	10.7	10.9	11.0	11.0	11.1	11.2	11.2	11.5
11.9	12.4	12.5	12.9	13.0	13.1	13.3	13.6	13.7	13.9
14.1	15.4	15.4	17.3	17.3	18.1	18.2	18.4	18.9	19.0
19.9	20.6	21.3	21.4	21.9	23.0	27.0	31.6	33.1	38.5

The descriptive statistics of the data are shown in Table 7 below.

Table 7. *Descriptive Statistics of the second data.*

N	Mean	Median	SD	Variance	Skewness	Kurtosis	Minimum	Maximum
100	9.777	7.85	7.2638	52.7634	1.4883	5.5485	0.8	38.5

Table 8. *MLEs (standard errors in parentheses) and the measures AIC, BIC, HQIC and CAIC of the first data.*

Distribution	$\hat{\mu}$	$\hat{\sigma}$	$\hat{a}$	$\hat{\lambda}$	AIC	BIC	HQIC	CAIC
TPG	0.3531289 (2.3676E-6)	0.6366211 (5.8268E-6)	1.2984443 (4.161E-11)	0.1645854 (6.0687E-7)	1001.7795	1007.5299	1004.1161	1001.8732
PG	1.0998154 (0.0000257)	1.1401995 (0.0000791)	0.1558073 (0.000012)	- -	1075.7277	1084.3533	1079.2327	1075.9167
TG	0.248029 (0.4844093)	1.3005865 (0.5463889)	- -	1.5884808 (0.396437)	1093.2296	1104.7304	1097.9029	1093.5471
TP	- -	- -	0.2037791 (0.0306841)	0.1642718 (0.1414493)	1060.4862	1069.1118	1063.9912	1060.6752
G	18.74255 (505.54667)	11.159707 (34.639565)	- -	- -	1185.1877	1196.6884	1189.8609	1185.5051
P	34.14885 (186.34564)	- -	- -	- -	1181.50961	1187.26001	1183.84626	1181.60336

Based on the Tables 5 and 8, we conclude that the new TPG model provides adequate fits as compared to other models in both applications with small values for SE, AIC, BIC, HQIC and CAIC. In the two applications, the proposed TPG model is much better than the five models.

7. Conclusions

In this paper we have proposed a new distribution, referred to as the TPG distribution. A mathematical treatment of the proposed distribution including explicit formulas for the density and survivor functions, moments, order statistics, and mean and median deviations have been provided. The estimation of the parameters has been approached by maximum likelihood. Also, the asymptotic variance-covariance matrix of the estimates has been obtained. The confidence intervals of the parameters of the new model are evaluated through the simulation study. Finally, two applications to real data indicate that the TPG distribution provides a good fit and can be used as a competitive model to fit real data.

Abbreviations

AIC	Akaike's Information Criterion
BIC	Bayesian Information Criterion
CAIC	Consistent Akaikes Information Criterion
HQIC	Hannan-Quinn Information Criterion
G	Gumbel
P	Power
PG	Power Gumbel
SE	Standard Error
TG	Transmuted Gumbel
TP	Transmuted Power
TPG	transmuted Power Gumbel

Author Contributions

Ahmed Ali Hurairah: Software, Supervision, Writing – original draft, Writing – review & editing

Nasr Tawfiq Almazaqi: Data curation, Formal Analysis, Investigation, Methodology, Resources

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] Abdul-Moniem, I. and Seham, M. Transmuted Gompertz Distribution. Computational and Applied Mathematics Journal, 2015, 1(3), 88-96.

[2] Afify, A. Z., Nofal, M., and Butt, N. S. Transmuted complementary Weibull geometric distribution, Pak. j. stat.oper.res. 2014, 10(4), 435-454.

[3] Akinsete, A., Famoye, F. and Lee, C. The beta-Pareto distribution. Statistics. 2008, 42, 547-563.

[4] Aryal, G. R. and Tsokos, C. P. Transmuted Weibull distribution: A generalization of the Weibull probability distribution. European Journal of Pure and Applied Mathematics. 2011, 4(2), 89-102.

[5] Azzalini, A. and Capitanio, A. Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t distribution. Journal of the Royal Statistical Society B. 2003, 65, 367-389.

[6] Barreto-Souza, W., Santos, A. H. S. and Cordeiro, G. M. The beta generalized exponential distribution. Journal of Statistical Computation and Simulation. 2010, 80, 159-172.

[7] Cordeiro, G. M., Gomez, A. E., De Silva, C. Q. and Ortega, E. E. M. The beta exponentiated Weibull distribution. Journal of Statistics Computation and Simulation. 2013, 83(1), 114-138.

[8] Cordeiro, G. M., Pescim, R. R. and Ortega, E. M. M. The Kumaraswamy Generalized Half-Normal Distribution for Skewed Positive Data. Journal of Data Science. 2012, 10(2012), 195-224.

[9] Cordeiro, G. M., Rodrigues, J. and de Castro, M. The exponential COM-Poisson distribution. Statistical Papers-STAT PAP. 2012, 53, 653-664.

[10] Cordeiro, G. M. and de Castro, M. A new family of generalized distributions. Journal of Statistical Computation and Simulation. 2011, https://doi.org/10.1080/0094965YY

[11] Cordeiro, G. M., Ortega, E. M. M. and Nadarajah, S. The Kumaraswamy Weibull distribution with application to failure data. Journal of the Franklin Institute. 2010, 347, 1399-1429.

[12] Elbatal, I. Transmuted modified inverse Weibull distribution: A generalization of the modified inverse Weibull probability distribution. International Journal of Mathematical Archive. 2013, 4(8), 117-129.

[13] Elbatal, I. and Aryal, G. On the transmuted additive Weibull distribution. Austrian Journal of Statistics. 2013, 42, 117-132, https://doi.org/10.17713/ajs.v42i2.160

[14] Eugene, N., Lee, C. and Famoye, F. Beta-normal distribution and its applications. Communications in Statistics, Theory and Methods. 2002, 31, 497-512.

[15] Ghitany, M. E., Al-awadhi, F. A., Alkhalfan, L. A. Marshall-Olkin extended Lomax Distribution and its application to censored Data, Communication in Statistics-Theory and Methods. 2007, 36, 1855–1866.

[16] Gupta, R. D. and Kundu, D. Exponentiated exponential family; an alternative to gamma and Weibull. Biometrical Journal. 2001, 43, 117 - 130.

[17] Hansen, B. E. Autoregressive conditional density estimation. International Economic Review. 1994, 35, 705-730.

[18] Khan, M. S. and King, R. Transmuted modified Weibull distribution: A generalization of the modified Weibull probability distribution. European Journal of Pure and Applied Mathematics. 2013, 6(1), 66-88.

[19] Kotz, S. & Nadarajah, S. Extreme Value Distributions: Theory and Applications, Imperial College Press, London, 2000.

[20] Nadarajah, S. The exponentiated Gumbel distribution with climate application Environmetrics. 2006, 17, 13–23.

[21] Nadarajah, S. and Eljabri, S. The Kumaraswami GP distribution. Journal of Data Science. 2013, 11, 739-766.

[22] Nadarajah, S. and Kotz, S. The beta Gumbel distribution. Math. Probab. Eng. 2004, 10, 323-332.

[23] Nadarajah, S. and Kotz, S. The beta exponential distribution. Reliability Engineering and System Safety. 2005, 91, 689-697.

[24] Shaw, W. T. and Buckley, I. R. C. The Alchemy of Probability Distributions: Beyond Gram Charlier Expansions, and a Skew-Kurtotic Normal Distribution from a Rank Transmutation Map. Technical Report No. 0901.0434, Cornell University, Ithaca, NY. 2007.

[25] Zhu, D. and Galbraith, J. W. A generalized asymmetric Student-t distribution with application to financial econometrics. Journal of Econometrics. 2010, 157, 297-305.

[26] Zografos, K. and Balakrishnan, N. On families of beta- and generalized gamma generated distributions and associated inference. Statistical Methodology. 2009, 6, 344-362.