mathematics

MDPI

Article

A Generalized Family of Exponentiated Composite Distributions

Bowen Liu † and **Malwane M. A. Ananda** *,†

Department of Mathematical Sciences, University of Nevada, Las Vegas, NV 89154, USA; bowen.liu@unlv.edu
* Correspondence: malwane.ananda@unlv.edu
† These authors contributed equally to this work.

Abstract: In this paper, we propose a new family of distributions, by exponentiating the random variables associated with the probability density functions of composite distributions. We also derive some mathematical properties of this new family of distributions, including the moments and the limited moments. Specifically, two special models in this family are discussed. Three real datasets were chosen, to assess the performance of these two special exponentiated-composite models. When fitting to these three datasets, these three special exponentiated-composite distributions demonstrate significantly better performance, compared to the original composite distributions.

Keywords: composite models; goodness-of-fit; IG distribution; Weibull distribution; Pareto distribution; exponential distributions; exponentiated models

MSC: 62P05; 62E99

Citation: Liu, B.; Ananda, M.M.A. A Generalized Family of Exponentiated Composite Distributions. *Mathematics* **2022**, *10*, 1895. https://doi.org/10.3390/math10111895

Academic Editors: Sławomir Nowaczyk, Rita P. Ribeiro and Grzegorz Nalepa

Received: 1 April 2022
Accepted: 29 May 2022
Published: 1 June 2022

1. Introduction

When dealing with right-skewed data, with a hump-shaped frequency distribution, analysts, generally, begin with fitting to a commonly used right-skewed parametric distribution. However, such a procedure, sometimes, fails to fit the data in a satisfactory way. For example, lognormal distribution is a natural choice, when fitting to right-skewed data. However, it does not provide satisfactory performance, when the upper tail is fairly large, while Pareto distribution is more suitable in such situations. If Pareto distribution is chosen instead, the lower tail features of the data cannot be captured, since the density of Pareto distribution is, monotonically, decreasing. Thus, the concept of composite distribution has arisen in the literature. The composite distributions are constructed, by combining two parametric distributions at a specific threshold.

Such a concept is widely used in different areas, such as modeling insurance claim size data [1–9], predicting the risk measures in insurance data analysis [4–7], fitting survival time data [10], and modeling precipitation data [11]. Such a concept demonstrates impressive performances, when the data is characterized with a very heavy upper tail, in which common distributions, such as normal or exponential distributions, cannot capture all of the data features. Due to the simplicity and the applicability of the concept, the researchers developed a considerable number of composite distributions, including Lognormal-Pareto [1], Weibull-Pareto [12], Weibull-Inverse Weibull [10], and so on.

The composite distributions seems proper, when modeling the data with heavy tails. For example, both the one-parameter Inverse Gamma- Pareto (IG-Pareto) model [6] and the one-parameter exponential-Pareto (exp-Pareto) model [13] were suggested as possible models, for insurance data modeling. However, they still cannot provide a satisfactory performance, when fitting to well-known insurance datasets, such as the Danish fire insurance dataset. Thus, it is necessary to improve the model. In order to improve the one-parameter IG-Pareto model, Liu and Ananda [8] proposed an exponentiated IG-Pareto model, by exponentiating the random variable associated with the probability density function (pdf) of an Inverse Gamma-Pareto distribution. With different datasets, their proposed model demonstrated significant improvement from the original model.

In fact, Liu and Ananda [8] suggested the idea of the exponentiated IG-Pareto model can be generalized, to all the existing composite distributions. Thus, in this paper, we propose a generalized family of exponentiated-composite distributions, with the derivation of some mathematical properties of this family.

The rest of the paper is organized as follows. Section 2 provides the formulation of the generalized exponentiated-composite distributions, as well as some of the mathematical properties of this new family of distributions. We, then, briefly, discuss two special exponentiated-composite distributions (exponentiated IG-Pareto and exponentiated exp-Pareto) in Section 3. In Section 4, we introduce a parameter-estimation method for exponentiated-composite models. The simulation results for the estimation method are, also, provided in Section 4. Three numerical examples are presented in Section 5. To conclude, a discussion is provided in Section 6.

2. Proposed Model: A Generalized Family of Exponentiated Composite Distributions

2.1. Model Formulation

Suppose X is a random variable that takes on non-negative real numbers. Let $f_X(x)$ be the probability density function (pdf) of X. A composite pdf $f_X(x)$ is defined [2] as follows:

$$f_X(x|\alpha_1,\alpha_2,\theta) = \begin{cases} cf_1(x|\alpha_1,\theta) & \text{if } x \in [0,\theta) \\ cf_2(x|\alpha_2,\theta) & \text{if } x \in [\theta,\infty) \end{cases}, \tag{1}$$

where c represents the normalizing constant, f_1 is the pdf of X, if X takes a value that is between 0 and θ, and α_1 represents the parameters of f_1; f_2 is the pdf of X, if X takes a value that is greater than θ, and α_2 represents the parameters of f_2. In real practice, we assume both f_1 and f_2 are smooth functions on their supports. However, the definition in (1) does not guarantee that $f_X(x)$ is a continuous differentiable function. To define a continuous differentiable pdf for composite distributions, the continuity and differentiability conditions are taken into account, as follows:

$$\begin{cases} \lim_{x\to\theta^-} f_X(x|\alpha_1,\alpha_2,\theta) = \lim_{x\to\theta^+} f_X(x|\alpha_1,\alpha_2,\theta) \\ \lim_{x\to\theta^-} \frac{df_X(x|\alpha_1,\alpha_2,\theta)}{dx} = \lim_{x\to\theta^+} \frac{df_X(x|\alpha_1,\alpha_2,\theta)}{dx}. \end{cases}$$

Essentially, the conditions can be summarized in a simpler way as: $f_1(\theta) = f_2(\theta)$ and $f_1'(\theta) = f_2'(\theta)$.

We apply a power transformation to the random variable X, by letting $Y = g(X) = X^{1/\eta}$, where g is monotone increasing for any $\eta > 0$. The inverse function $g^{-1}(Y) = Y^\eta$ has continuous derivative on $(0,\infty)$, for any $\eta > 0$. Then, the pdf of Y is given by:

$$f_Y(y|\alpha_1,\alpha_2,\theta,\eta) = \begin{cases} cf_1(y^\eta|\alpha_1,\theta)\eta y^{\eta-1} & \text{if } y \in [0,\theta^{\frac{1}{\eta}}) \\ cf_2(y^\eta|\alpha_2,\theta)\eta y^{\eta-1} & \text{if } y \in [\theta^{\frac{1}{\eta}},\infty). \end{cases} \tag{2}$$

We shall prove that the pdf of Y is, still, a continuous differentiable pdf of a composite distribution.

Theorem 1. *If X is a random variable associated with a continuous differentiable pdf of a composite distribution, then $Y = X^{\frac{1}{\eta}}$, also, has a continuous differentiable pdf of a composite distribution for all $\eta > 0$.*

Proof of Theorem 1. Let $u = \theta^{\frac{1}{\eta}}$.

We first show that $\lim_{y\to u^-} f_Y(y|\alpha_1,\theta,\eta) = \lim_{y\to u^+} f_Y(y|\alpha_2,\theta,\eta)$.

From (2), we have

$$
\begin{aligned}
\lim_{y\to u^-} f_Y(y|\alpha_1,\theta,\eta) &= cf_1(u^\eta|\alpha_1,\theta)\eta u^{\eta-1} \\
&= cf_1(\theta|\alpha_1,\theta)\eta u^{\eta-1} \\
&= cf_2(\theta|\alpha_2,\theta)\eta u^{\eta-1} \\
&= cf_2(u^\eta|\alpha_2,\theta)\eta u^{\eta-1} \\
&= \lim_{y\to u^+} f_Y(y|\alpha_2,\theta,\eta).
\end{aligned}
$$

Then, we want to show that $\lim_{y\to u^-} \frac{df_Y(y|\alpha_1,\theta,\eta)}{dy} = \lim_{y\to u^+} \frac{df_Y(y|\alpha_1,\theta,\eta)}{dy}$.
By the chain rules, we have:

$$
\begin{aligned}
\lim_{y\to u^-} \frac{df_Y(y|\alpha_1,\alpha_2,\theta,\eta)}{dy} &= c(f_1^{'}(u^\eta|\alpha_1,\theta)\eta u^{\eta-1} + f_1(u^\eta|\alpha_1,\theta)(\eta-1)u^{\eta-2}) \\
&= c(f_2^{'}(u^\eta|\alpha_2,\theta)\eta u^{\eta-1} + f_2(u^\eta|\alpha_2,\theta)(\eta-1)u^{\eta-2}) \\
&= \lim_{y\to u^+} \frac{df_Y(y|\alpha_1,\alpha_2,\theta,\eta)}{dy}.
\end{aligned}
$$

Thus, $f_Y(y|\alpha_1,\alpha_2,\theta,\eta)$ is a continuous differentiable pdf of a composite distribution. □

Since $f_Y(y)$ is still a composite distribution, we hereby name $f_Y(y)$ as **the exponentiated-composite distribution induced by the parent composite distribution** $f_X(x)$. Correspondingly, we, also, name Y as **the exponentiated-composite random variable induced by the parent composite random variable** X.

Property 1. *If Y is a random variable associated with a pdf of an exponentiated-composite distribution induced by X, where X is associated with a pdf of a continuous differentiable composite distribution, then any exponetiated composite distribution induced by Y is induced by X.*

Proof of Property 1. The proof is simple. Let $U = Y^{\frac{1}{\gamma}}$

Since Y is induced by X, there exists $\eta > 0$, such that $Y = X^{\frac{1}{\eta}}$.

Hence, $U = Y^{\frac{1}{\gamma}} = (X^{\frac{1}{\eta}})^{\frac{1}{\gamma}} = X^{\frac{1}{\eta\gamma}}$. Therefore, U is induced by X. □

Notice that property 1, essentially, gives us the general idea that if a random variable Y is induced by X, where X is a random variable associated with a composite distribution, then exponentiating Y, eventually, leads to a distribution of the same type as Y. For example, a Weibull-Inverse Weibull composite [10] can be seen as a distribution induced by an Exponential-Inverse Exponential composite. Hence, by exponentiating a Weibull-Inverse Weibull composite random variable, we, still, obtain a Weibull-Inverse Weibull composite random variable.

2.2. Mathematical Properties

In this subsection, we derive the moments and the limited moments of Y, as an induced exponentiated-composite random variable of the composite random variable X. Before any derivations, we need to guarantee that Y has finite moments of all orders, if X has finite moments of all orders.

Property 2. *Suppose X is a composite random variable. If X has finite moments of all orders, then the exponentiated-composite random variable Y induced by X, also, has finite moments of all orders.*

Proof of Property 2. Suppose X has finite moments of all orders. Then, $E(X^p) < \infty$, for all $p \in \mathbb{R}$. Since, Y is induced by X, $Y = X^{\frac{1}{\eta}}$. Then, for any $p \in \mathbb{R}$, $E(Y^p) = E(X^{\frac{p}{\eta}}) < \infty$. Thus, Y has finite moments of all orders, if X has finite moments of all orders. □

The following derivations are based on the assumption that X has finite moments of all orders. Assume the k-th moment of X is μ_k.

2.2.1. *k*-th Moment

Since we defined that $Y = X^{\frac{1}{\eta}}$, the k-th moment of Y can be simply derived as follows:

$$E(Y^k) = E(X^{\frac{k}{\eta}}) = \mu_{\frac{k}{\eta}}. \tag{3}$$

Essentially, the moment of Y is a real-valued fractional moment of X.

2.2.2. Limited *k*-th Moment

Limited moment is a widely used concept in the insurance industry. Suppose X represents the loss value reported in a claim. Let b be the maximum value that an insurance company can pay, under a specific insurance policy. That is, when the reported loss is greater than b, the company will pay the amount b, instead of the actual reported loss [14]. Formally, the limited-loss random variable $X \wedge b$ is defined, as follows, given $b > 0$:

$$X \wedge b = \begin{cases} x & \text{if } x < b \\ b & \text{if } x \geq b. \end{cases}$$

Respectively, $E[(X \wedge b)^t]$, as the tth limited moment of a random variable X, can be defined as:

$$E[(X \wedge b)^t] = \int_0^b x^t f_X(x)dx + \int_b^\infty b^t f_X(x)dx.$$

It follows that if X is L_p integrable, then the t-th limited moment of X exists.

For any valid probability distribution, defined on $\mathcal{R}$, the cumulative distribution function (CDF) must exist. Thus, without loss of generality, let the CDF of X be $F_X(x)$ and define the incomplete moment function as $M(u;r) = \int_0^u x^r f_X(x)dx$. Then, the t-th limited moment of X can be rewritten, as follows:

$$E[(X \wedge b)^t] = M(b;t) + b^t[1 - F_X(b)].$$

Suppose $Y = X^{\frac{1}{\eta}}$. Then, the corresponding pdf of Y is $f_Y(y) = f_X(y^\eta)\eta y^{\eta-1}$. Therefore, we can explicitly represent the t-th limited moment of Y as follows, with the incomplete moment function of X and the CDF of X:

$$\begin{aligned} E[(Y \wedge b)^t] &= \int_0^b y^t f_Y(y)dy + \int_b^\infty b^t f_Y(y)dy \\ &= \int_0^b y^t f_X(y^\eta)\eta y^{\eta-1}dy + \int_b^\infty b^t f_X(y^\eta)\eta y^{\eta-1}dy \\ &= \int_0^{b^\eta} x^{\frac{t}{\eta}} f_X(x)dx + \int_{b^\eta}^\infty b^t f_X(x)dx \\ &= M(b^\eta; \frac{t}{\eta}) + b^t[1 - F_X(b^\eta)]. \end{aligned} \tag{4}$$

Moreover, assume that X is a random variable associated with a continuous differentiable pdf, with a form in (1). Define the following:

$$\begin{aligned} M_1(u;r) &= \int_0^u x^r f_1(x)dx, \\ M_2(u;r) &= \int_0^u x^r f_2(x)dx, \\ F_1(u) &= \int_0^u f_1(x)dx, \\ F_2(u) &= \int_0^u f_2(x)dx. \end{aligned}$$

Then $E[(X \wedge b)^t]$ can be expressed, explicitly, with the above quantities, as follows:

$$E[(X \wedge b)^t] = \begin{cases} cM_1(b;t) + cb^t[F_1(\theta) - F_1(b)] + cb^t(1 - F_2(\theta)) & \text{if } b \in (0, \theta) \\ cM_1(b;t) + cb^t[1 - F_2(b)] & \text{if } b = \theta \\ cM_1(\theta;t) + cM_2(b;t) - cM_2(\theta;t) + cb^t[1 - F_2(b)] & \text{if } b \in (\theta, \infty), \end{cases}$$

Suppose Y is an exponentiated-composite random variable induced by X, where the general form of Y is defined in (2). We can, also, explicitly, represent $E[(Y \wedge b)^t]$, as follows:

$$E[(Y \wedge b)^t] = \begin{cases} cM_1(b^\eta; \frac{t}{\eta}) + cb^t[F_1(\theta) - F_1(b^\eta)] + cb^t[1 - F_2(\theta)] & \text{if } b \in (0, \theta^{\frac{1}{\eta}}) \\ cM_1(b^\eta; \frac{t}{\eta}) + cb^t[1 - F_2(\theta)] & \text{if } b = \theta^{\frac{1}{\eta}} \\ cM_1(\theta; \frac{t}{\eta}) + cM_2(b^\eta; \frac{t}{\eta}) - cM_2(\theta; \frac{t}{\eta}) + cb^t[1 - F_2(b^\eta)] & \text{if } b \in (\theta^{\frac{1}{\eta}}, \infty). \end{cases} \tag{5}$$

3. Special Distributions

3.1. Two-Parameter Exponentiated Inverse Gamma-Pareto Model

The one-parameter composite IG-Pareto model was introduced by Aminzadeh and Deng [6]. Suppose a random variable X follows a one-parameter composite Inverse Gamma-Pareto distribution, such that the pdf of X is as follows:

$$f_X(x|\theta) = \begin{cases} \frac{c(k\theta)^\alpha x^{-\alpha-1} e^{-\frac{k\theta}{x}}}{\Gamma(\alpha)} & \text{if } x \in [0, \theta) \\ \frac{c(\alpha-k)\theta^{\alpha-k}}{x^{\alpha-k+1}} & \text{if } x \in [\theta, \infty), \end{cases}$$

where $c = 0.711384, k = 0.144351, \alpha = 0.308298$. The constants c, k, and α are obtained, by imposing the continuity and differentiability conditions mentioned in Section 2.1.

By utilizing (2), Liu and Ananda [8] developed the two-parameter exponentiated Inverse Gamma-Pareto model. The model has the pdf, as follows:

$$f_Y(y|\theta, \eta) = \begin{cases} \frac{c(k\theta)^\alpha (y^\eta)^{-\alpha-1} e^{-\frac{k\theta}{y^\eta}}}{\Gamma(\alpha)} \eta y^{\eta-1} & \text{if } y \in [0, \theta^{\frac{1}{\eta}}) \\ \frac{c(\alpha-k)\theta^{\alpha-k}}{(y^\eta)^{\alpha-k+1}} \eta y^{\eta-1} & \text{if } y \in [\theta^{\frac{1}{\eta}}, \infty). \end{cases} \tag{6}$$

The t-th raw moment can be easily derived, as follows:

$$E(Y^t) = c\left[\frac{(k\theta)^{\frac{t}{\eta}}}{\Gamma(\alpha)} \Gamma(\alpha - \frac{t}{\eta}, k) - \frac{(\alpha - k)\theta^{k-\alpha+\frac{t}{\eta}}}{k - \alpha + \frac{t}{\eta}}\right], \tag{7}$$

where $\Gamma(.,.)$ stands for an upper incomplete gamma function, $\Gamma(\alpha, x) = \int_x^\infty t^{\alpha-1} e^{-t} dt$. When $k - \alpha + \frac{t}{\eta} < 0$, the t-th moment is finite. The t-th limited moment of Y was derived by Liu and Ananda [8], as follows:

$$E[(Y \wedge b)^t] = \begin{cases} c\left[\frac{\Gamma(\alpha - \frac{t}{\eta}, \frac{k\theta}{b^\eta})(k\theta)^{\frac{t}{\eta}} + b^t\Gamma(\alpha,k) - b^t\Gamma(\alpha, \frac{k\theta}{b^\eta})}{\Gamma(\alpha)} + b^t\right] & \text{if } b \in (0, \theta^{1/\eta}) \\ c\left[\frac{\Gamma(\alpha - \frac{t}{\eta}, k)(k\theta)^{\frac{t}{\eta}}}{\Gamma(\alpha)} + b^t\right] & \text{if } b = \theta^{1/\eta} \\ c\left\{\frac{\Gamma(\alpha - \frac{t}{\eta}, k)(k\theta)^{\frac{t}{\eta}}}{\Gamma(\alpha)} + \frac{(\alpha-k)[b^{t-\eta(\alpha-k)}\theta^{\alpha-k} - \theta^{\frac{t}{\eta}}]}{k - \alpha + \frac{t}{\eta}} + b^{t-(\alpha-k)\eta}\theta^{\alpha-k}\right\} & \text{if } b \in (\theta^{1/\eta}, \infty). \end{cases}$$

Figure 1 is the plot of exponentiated IG-Pareto distributions, for different values of η and θ. The extra-exponent parameter η introduces additional flexibility to the original model. When $\eta \geq 1$, the distribution is associated with a unimodal pdf.

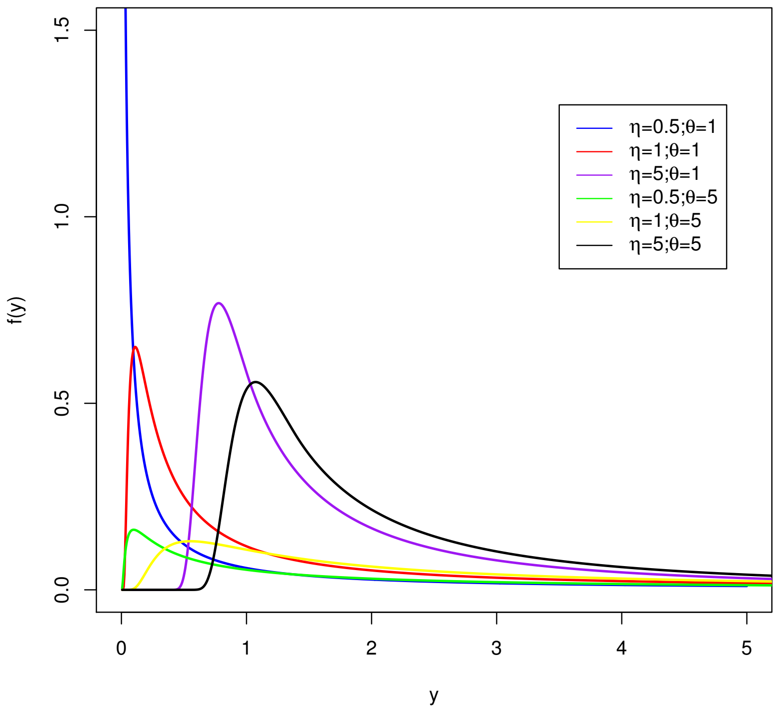

Figure 1. Plots of the exponentiated IG-Pareto density function, for some parameter values.

3.2. Two-Parameter Exponentiated Exp-Pareto Model

The one-parameter composite exp-Pareto model was introduced by Teodorescu [13]. Suppose a random variable X follows a one-parameter composite exp-Pareto distribution, such that the pdf of X is as follows:

$$f_X(x|\theta) = \begin{cases} c(\frac{\alpha+1}{\theta})e^{-\frac{(\alpha+1)x}{\theta}} & \text{if } x \in [0,\theta) \\ c\alpha\frac{\theta^\alpha}{x^{\alpha+1}} & \text{if } x \in [\theta,\infty), \end{cases}$$

where $c = 0.574, \alpha = 0.349976$. The constants c and α are obtained, by imposing the continuity and differentiability conditions mentioned in Section 2.1. By utilizing (2), the corresponding two-parameter exponentiated exp-Pareto pdf is as follows:

$$f_Y(y|\theta,\eta) = \begin{cases} c(\frac{\alpha+1}{\theta})e^{-\frac{(\alpha+1)y^\eta}{\theta}}\eta y^{\eta-1} & \text{if } y \in [0,\theta^{\frac{1}{\eta}}) \\ c\alpha\frac{\theta^\alpha}{(y^\eta)^{\alpha+1}}\eta y^{\eta-1} & \text{if } y \in [\theta^{\frac{1}{\eta}},\infty). \end{cases} \tag{8}$$

The t-th raw moment of the exponentiated exp-Pareto is derived, as follows:

$$E(Y^t) = c(\frac{\theta}{\alpha+1})^{\frac{t}{\eta}}\left[\Gamma(\frac{t}{\eta}+1) - \Gamma(\frac{t}{\eta}+1,\alpha+1)\right] + \frac{c\alpha\theta^{\frac{t}{\eta}}}{\alpha-\frac{t}{\eta}}, \tag{9}$$

where $\Gamma(\alpha)$ represents a Gamma function and $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1}e^{-t}dt$. When $\frac{t}{\eta} < \alpha$, the t-th moment is finite. The t-th limited moment of the exponentiated exp-Pareto is:

$$E[(Y\wedge b)^t] = \begin{cases} c\{(\frac{\theta}{\alpha+1})^{\frac{t}{\eta}}[\Gamma(\frac{t}{\eta}+1) - \Gamma(\frac{t}{\eta}+1,\frac{\alpha+1}{\theta}b^\eta)] + b^t(e^{-\frac{\alpha+1}{\theta}b^\eta} - e^{-\alpha-1}) + b^t\} & \text{if } b < \theta^{1/\eta} \\ c\{(\frac{\theta}{\alpha+1})^{\frac{t}{\eta}}[\Gamma(\frac{t}{\eta}+1) - \Gamma(\frac{t}{\eta}+1,\alpha+1)] + b^t\} & \text{if } b = \theta^{1/\eta} \\ c\{(\frac{\theta}{\alpha+1})^{\frac{t}{\eta}}[\Gamma(\frac{t}{\eta}+1) - \Gamma(\frac{t}{\eta}+1,\alpha+1)] + \frac{c\alpha\theta^\alpha}{\frac{t}{\eta}-\alpha}(b^{t-\frac{\alpha}{\eta}} - \theta^{\frac{t}{\eta}-\alpha}) + b^t\} & \text{if } b > \theta^{1/\eta}. \end{cases}$$

Figure 2 is the plot of the exponentiated exp-Pareto distributions, for different values of η and θ. When $\eta > 1$, the pdf is unimodal. For the case that $\eta < 1$, the pdf is, monotonically, decreasing.

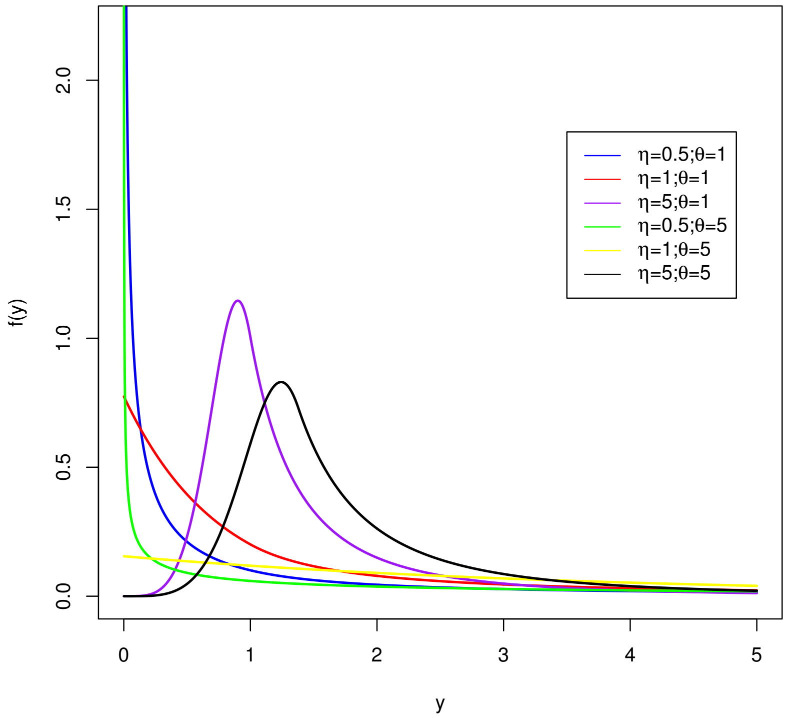

Figure 2. Plots of the exponentiated exp-Pareto density function, for some parameter values.

4. Parameter Estimation and Simulation Studies

4.1. Parameter Search Method

In this section, we want to, briefly, explain the parameter estimation of the two-parameter exponentiated-composite models. Essentially, since the estimation of the location parameter is involved in the composite modeling, a special algorithm for the composite-model-parameter estimation was developed and used [1,6,7]. For some special exponentiated-composite models, such an algorithm could be adapted to estimate the parameters. Liu and Ananda [8] showed that the parameter estimation of the exponentiated IG-Pareto model can be accomplished, using a step-wise grid-search procedure on the exponent parameter η. Suppose $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)$ is an independent and identically distributed (i.i.d.) sample from a two-parameter exponentiated-composite distribution, with a pdf defined in (2). Let $L(\boldsymbol{y}|\theta, \eta)$ be the likelihood of $\boldsymbol{y}$. Assuming within the sample $\boldsymbol{y}$, let m observations be less than $\theta^{\frac{1}{\eta}}$. Now, suppose for every fixed η, the analytical solution of $\frac{\partial L(\boldsymbol{y}|\theta,\eta)}{\partial \theta} = 0$ exists, where θ could be represented as a function of η, m, and $\boldsymbol{y}$. Denote the function as $\theta(\eta, m, \boldsymbol{y})$.

I. Arrange the observations in the sample in an increasing order, such that $y_{(1)} \leq y_{(2)} \leq \ldots \leq y_{(n)}$.
II. Define a candidate set $\boldsymbol{\eta}_c$, for the exponent parameter η. For example, we can choose $\boldsymbol{\eta}_c$ as an increasing sequence that starts at 0.01 and ends at 20, with an increment by 0.01.
III. For each $\eta \in \boldsymbol{\eta}_c$, we start with $m = 1$ (assuming the $m = 1$ observation is below $\theta^{\frac{1}{\eta}}$). Calculate the estimate of θ, based on $\theta(\eta, m, \boldsymbol{y})$. Denote this estimate as $\hat{\theta}_{1,\eta}$. If $y_1^{\eta} \leq \hat{\theta}_{1,\eta} \leq y_2^{\eta}$, then $m = 1$. If not, go to step IV.
IV. Let $m = 2$. Calculate the estimate of θ, based on $\theta(\eta, m, \boldsymbol{y})$. Denote this estimate as $\hat{\theta}_{2,\eta}$. If $y_2^{\eta} \leq \hat{\theta}_{2,\eta} \leq y_3^{\eta}$, then $m = 2$. The above steps will resume, until m is detected. Once m is detected, keep $\hat{\theta}_{m,\eta}$ as the estimate of θ, for the corresponding η.

V. The estimate of η is, then, identified as the $\hat{\eta}$ that maximizes the likelihood function: $L(\boldsymbol{y}|\theta,\eta)$:

$$\hat{\eta} = \arg\max_{\eta \in \boldsymbol{\eta}_c} L(\boldsymbol{y}|\theta,\eta)$$

The corresponding estimate of θ is, then, determined as $\theta(\hat{\eta}, m, \boldsymbol{y})$.

With the above algorithm, the accuracy of the estimates from the two-parameter exponentiated IG-Pareto model was demonstrated, with limited simulations. Thus, in the next section, we demonstrate that the parameter estimation of the two-parameter exponentiated exp-Pareto distribution can be done, based on the above procedure.

4.2. Estimation of the Parameters of the Exponentiated Exp-Pareto Model

To utilize the search procedure in Section 4.1, the closed-form solution of $\frac{\partial L(\boldsymbol{y}|\theta,\eta)}{\partial\theta} = 0$ needs to be obtained. Suppose $y_1, y_2, \ldots, y_n$ are i.i.d. two-parameter exponentiated exp-Pareto random variables, with the pdf defined in (8), and assume $y_1 < y_2 < \ldots < y_n$, without loss of generality. Assume that there exists a value $m \in \{1, 2, \ldots, n-1\}$, such that $y_m^\eta < \theta < y_{m+1}^\eta$. The likelihood $L(\boldsymbol{y}|\theta,\eta)$ could be written, as follows:

$$\begin{aligned} L(\boldsymbol{y}|\theta,\eta) &= \prod_{i=1}^{m} c\left(\frac{\alpha+1}{\theta}\right)e^{-\frac{(\alpha+1)y_i^\eta}{\theta}}\eta y_i^{\eta-1} \prod_{j=m+1}^{n} c\alpha\frac{\theta^\alpha}{(y_j^\eta)^{\alpha+1}}\eta y_j^{\eta-1} \\ &= \frac{c^n\eta^n(\alpha+1)^m\alpha^{n-m}(\prod_{i=1}^m y_i)^{\eta-1}}{(\prod_{j=m+1}^n y_j)^{\alpha\eta+1}}\theta^{\alpha n-(\alpha+1)m}e^{-\frac{\alpha+1}{\theta}\sum_{i=1}^m y_i^\eta}, \end{aligned}$$

Correspondingly, the log-likelihood $l(\boldsymbol{y}|\theta,\eta)$ could be written as:

$$\begin{aligned} l(\boldsymbol{y}|\theta,\eta) = n\log(c) + n\log(\eta) + m\log(\alpha+1) + (n-m)\log(\alpha) + (\eta-1)\sum_{i=1}^{m}\log(y_i) \\ - (\alpha\eta+1)\sum_{j=m+1}^{n}\log(y_j) + [\alpha n - (\alpha+1)m]\log(\theta) - \frac{\alpha+1}{\theta}\sum_{i=1}^{m} y_i^\eta \end{aligned} \tag{10}$$

Therefore, given η and m, the closed-form solution of $\frac{\partial l(\boldsymbol{y}|\theta,\eta)}{\partial\theta} = 0$ can be obtained as:

$$\theta(\eta,\boldsymbol{y}) = \frac{(\alpha+1)\sum_{i=1}^m y_i^\eta}{(\alpha+1)m - \alpha n}. \tag{11}$$

Therefore, we can utilize the search method in Section 4.1, to estimate the parameters of an exponentiated exp-Pareto model.

4.3. The Estimation of Standard Error with the Observed Fisher Information Matrix

Let $\Theta = (\theta, \eta)$ be the parameters of an exponentiated exp-Pareto pdf, defined in (8). Suppose $\hat{\Theta} = (\hat{\theta}, \hat{\eta})$ is the maximum likelihood estimate (MLE) of Θ. Notice that, with proper regularity conditions,

$$\mathcal{I}(\Theta)^{-\frac{1}{2}}(\hat{\Theta} - \Theta) \xrightarrow{\mathcal{D}} \mathcal{N}(\boldsymbol{0}_2, \boldsymbol{I}_2), \text{as } n \to \infty, \tag{12}$$

where $\mathcal{I}(\Theta)$ is the Fisher information matrix of Θ, $\mathcal{D}$ stands for convergence in distribution and $\mathcal{N}(\boldsymbol{0}_2, \boldsymbol{I}_2)$ is the bivariate standard normal distribution. Thus, the asymptotic variance of $\hat{\Theta}$ can be derived with the Fisher information matrix $\mathcal{I}(\Theta) = -E(\frac{\partial^2 l(\boldsymbol{\Theta})}{\partial\boldsymbol{\Theta}\partial\boldsymbol{\Theta}^T})$, where l stands for the the log-likelihood in (10). Define $I(\Theta) = -\frac{\partial^2 l(\boldsymbol{\Theta})}{\partial\boldsymbol{\Theta}\partial\boldsymbol{\Theta}^T}$. The expression of $I(\Theta)$ could be derived, explicitly, as follows:

$$I(\Theta) = \begin{pmatrix} -\frac{\partial^2 l(\boldsymbol{y}|\theta,\eta)}{\partial\theta^2} & -\frac{\partial^2 l(\boldsymbol{y}|\theta,\eta)}{\partial\theta\partial\eta} \\ -\frac{\partial^2 l(\boldsymbol{y}|\theta,\eta)}{\partial\theta\partial\eta} & -\frac{\partial^2 l(\boldsymbol{y}|\theta,\eta)}{\partial\eta^2} \end{pmatrix}$$
$$= \begin{pmatrix} \frac{\alpha n-(\alpha+1)m}{\theta^2} + \frac{2(\alpha+1)}{\theta^3}\sum_{i=1}^m y_i^\eta & -\frac{\alpha+1}{\theta^2}\sum_{i=1}^m y_i^\eta \log(y_i) \\ -\frac{\alpha+1}{\theta^2}\sum_{i=1}^m y_i^\eta \log(y_i) & \frac{n}{\eta^2} + \frac{\alpha+1}{\theta}\sum_{i=1}^m y_i^\eta [\log(y_i)]^2 \end{pmatrix}.$$

Theoretically, the asymptotic covariance matrix of Θ equals $E(\mathcal{I}(\Theta))$. However, the expectation of the Fisher information matrix cannot be evaluated in real applications. Therefore, the asymptotic variance of $\hat{\Theta}$ is approximated by $I(\hat{\Theta})^{-1}$, the inverse of the observed Fisher information matrix. Correspondingly, the standard error of $\hat{\theta}$ and $\hat{\eta}$ are estimated, by taking the square root of the diagonal elements of $I(\hat{\Theta})^{-1}$.

4.4. Simulations

To assess the accuracy for the estimates of $\hat{\theta}$ and $\hat{\eta}$, from the search method, a simulation study was done for the chosen sample size n, θ, and η. For each simulation scenario, $r = 2000$ samples were generated from the composite density, given in (8). The R package 'mistr' [15] was used to generate the random samples from the composite pdf, defined in (8).

Table 1 presents the results from all of the simulation scenarios. The means, the average standard errors (SE), and the standard deviations (SD) of the estimates are provided. The SE values of the estimates were obtained, using the observed Fisher information matrix given in Section 4.3. For all the simulation scenarios, we spotted that the average of the estimates of θ gets closer to the true θ, as the sample size n increases. This is, also, the case for the average of the estimates of η. Additionally, for both θ and η, the average SE and the SD of the estimates decrease, as the sample size n increases. As the sample size n increases, the accuracy of the estimates found by the algorithm becomes more accurate. This is expected, due to the consistency property of the maximum likelihood estimates.

Table 1. Simulation results for the parameter estimation of the exponentiated exp-Pareto distributions.

Scenarios			$n = 50$			$n = 100$			$n = 200$			$n = 500$		
η	θ	**Par.** [1]	**Mean**	**SD**	**SE**	**Mean**	**SD**	**SE**	**Mean**	**SD**	**SE**	**Mean**	**SD**	**SE**
0.8	1	$\hat{\eta}$	0.828	0.119	0.113	0.811	0.080	0.078	0.807	0.055	0.055	0.803	0.034	0.035
		$\hat{\theta}$	1.046	0.354	0.328	1.017	0.235	0.225	1.000	0.153	0.156	1.005	0.098	0.099
5	1	$\hat{\eta}$	5.158	0.749	0.702	5.077	0.506	0.489	5.042	0.354	0.343	5.011	0.210	0.216
		$\hat{\theta}$	1.046	0.347	0.326	1.018	0.242	0.226	1.007	0.158	0.157	1.002	0.099	0.099
0.8	5	$\hat{\eta}$	0.828	0.117	0.113	0.815	0.078	0.078	0.807	0.054	0.055	0.803	0.035	0.035
		$\hat{\theta}$	5.551	2.043	1.882	5.234	1.269	1.226	5.135	0.850	0.841	5.058	0.532	0.521
5	5	$\hat{\eta}$	5.140	0.714	0.700	5.072	0.505	0.488	5.047	0.351	0.343	5.023	0.208	0.216
		$\hat{\theta}$	5.503	1.997	1.860	5.259	1.289	1.230	5.125	0.859	0.840	5.064	0.517	0.522
10	5	$\hat{\eta}$	10.312	1.461	1.404	10.174	1.004	0.979	10.068	0.682	0.685	10.029	0.433	0.432
		$\hat{\theta}$	5.527	2.006	1.873	5.231	1.287	1.225	5.113	0.844	0.838	5.049	0.532	0.520
10	20	$\hat{\eta}$	10.324	1.406	1.406	10.158	0.990	0.978	10.119	0.703	0.689	10.023	0.430	0.431
		$\hat{\theta}$	24.183	13.175	11.360	21.995	7.403	7.021	21.112	4.896	4.684	20.261	2.848	2.801
0.4	2	$\hat{\eta}$	0.414	0.058	0.056	0.408	0.040	0.039	0.404	0.028	0.027	0.401	0.017	0.017
		$\hat{\theta}$	2.112	0.670	0.638	2.048	0.459	0.437	2.030	0.316	0.304	2.011	0.189	0.190
0.4	1	$\hat{\eta}$	0.415	0.059	0.057	0.407	0.041	0.039	0.402	0.028	0.027	0.402	0.018	0.017
		$\hat{\theta}$	1.033	0.353	0.325	1.018	0.231	0.225	1.009	0.159	0.158	1.001	0.097	0.099
4	5	$\hat{\eta}$	4.141	0.578	0.564	4.060	0.400	0.391	4.029	0.267	0.274	4.012	0.173	0.173
		$\hat{\theta}$	5.515	2.080	1.864	5.206	1.292	1.218	5.100	0.839	0.835	5.038	0.530	0.519
4	1	$\hat{\eta}$	4.125	0.576	0.562	4.052	0.381	0.390	4.035	0.277	0.275	4.016	0.171	0.173
		$\hat{\theta}$	1.042	0.339	0.326	1.021	0.233	0.226	1.012	0.159	0.158	1.000	0.098	0.099

[1] Par. stands for the estimated parameter.

5. Real Data Application

We applied the exponentiated IG-Pareto model and the exponentiated exp-Pareto model to the well-known Danish fire insurance data, Norwegian fire insurance data, and Society of Actuaries (SOA) group medical claims data. We compared these two models with the original one-parameter IG-Pareto and exp-Pareto models. For additional comparison purposes, we included the Weibull model, the Inverse Gamma model, and the Weibull-Inverse Weibull composite model. We chose the Weibull-Inverse Weibull composite model, since it demonstrates the best performance among all the composite models, when fitting to the Danish fire insurance data [5].

We used four different goodness-of-fit (GoF) measures, to compare the performances of the different models. The brief description of the measures are listed, as follows ($\boldsymbol{y}$ denotes the observed sample):

1. The negative log-likelihood (NLL) is defined, in the following manner:
$$NLL = -logL(\hat{\boldsymbol{\theta}}|\boldsymbol{y}).$$
2. The definition of the Akaike information criterion (AIC) [16] is provided, in the following manner:
$$AIC = -2logL(\hat{\boldsymbol{\theta}}|\boldsymbol{y}) + 2k,$$
where k denotes the number of model parameters.
3. The Bayesian information criterion (BIC) [16] is defined, as follows:
$$BIC = -2logL(\hat{\boldsymbol{\theta}}|\boldsymbol{y}) + klog(n),$$
where k stands for the number of model parameters and n represents the sample size.
4. Kolmogorov–Smirnoff (KS) statistic: The KS statistic is defined, as follows [17]:
$$D_n = \sup_y |F(\hat{\boldsymbol{\theta}}|\boldsymbol{y}) - F_n(\boldsymbol{y})|,$$
where $F_n(\boldsymbol{y})$ stands for the emipirical CDF of the observed sample $\boldsymbol{y}$.

In addition, value-at-risk (VaR) was used to evaluate the performances of the different models, when fitting to the upper tail of the data. The definition of VaR at the level of α is provided, as follows:
$$VaR_\alpha(Y) = F_Y^{-1}(\alpha),$$
where $F_Y^{-1}(.)$ is the inverse CDF function of the random variable Y. In Solvency II, designated by the European Union (EU), VaR is regarded as an important risk measur,e to determine the required amount that an insurance company needs to hold against the potential insolvency.

We obtained the estimates of the parameters in all the models, NLL, AIC, BIC, KS distance, and VaR, using the R software.

5.1. Danish Fire Insurance Data

To assess the performance of different models, the well-known Danish fire insurance data were used. This dataset was utilize,d by various statisticians, to investigate the performances of different loss models [1,2,5,7,8]. There are 2492 fire insurance claims reported in Denmark, from the years 1980 to 1990, in this dataset. Inflation adjustment was completed, for the claim sizes. The data have been scaled to one million Danish Krones (DKK), for analysis purposes. We accessed the dataset from the *SMPracticals* package in R software [18]. Figure 3 is the histogram of this dataset. The detailed summary of the dataset is provided in Table 2.

Table 2. Summary statistics, for Danish fire insurance data (in 1 million DKKs).

Sample Size	Mean	SD	Minimum	Q1	Q2	Q3	Maximum
2492	3.06	7.98	0.31	1.16	1.63	2.65	263.25

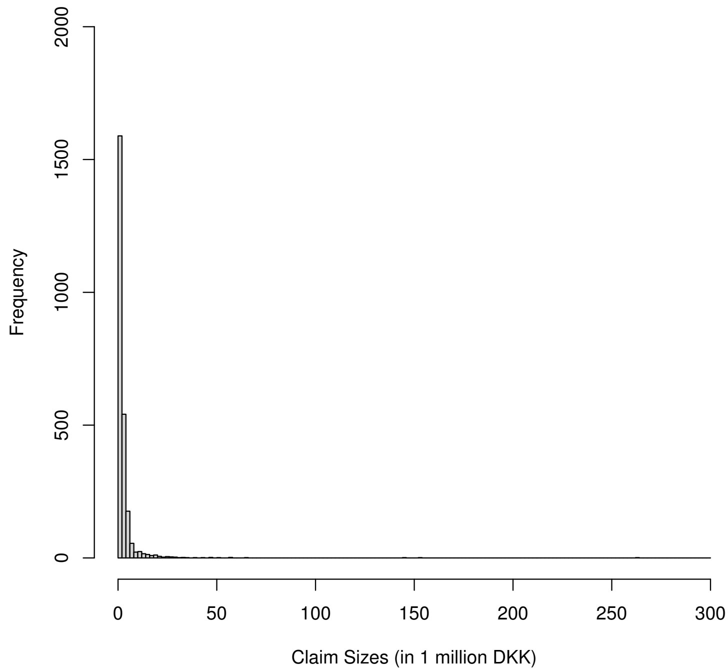

Figure 3. Histogram of Danish fire insurance dataset.

The model comparisons, in terms of four GoF measures, are presented in Table 3. The VaR estimates for all the models, at 0.90 and 0.95 levels are, also, provided. When fitting to the Danish Fire Insurance Data, the exponentiated IG-Pareto model demonstrated a significantly better performance, compared to the original one-parameter IG-Pareto model. Similarly, the exponentiated exp-Pareto model, also, presented a significant improvement from the exp-Pareto model, in terms of all measures. These results are in accordance with Figures 4 and 5. Figure 4 shows the comparison of the fitted composite models, along with the Gaussian kernel density estimate of the Danish fire insurance sata. Both exponentiated models provide a good fit to the Danish Fire Insurance Data, while the original one-parameter composite models perform poorly. This is observed in Figure 5. Figure 5 is the comparison of the CDF of the fitted models, with the empirical CDF of the Danish fire insurance data. The exponentiated-composite models provided a closer fit to the empirical CDF, compared to the original composite models. It is, also, noticeable that the two-parameter exponentiated exp-Pareto model performed better than the common distributions, such as Weibull and IG, in terms of all the measures. However, both the Weibull-Inverse Weibull model and the Weibull-Pareto model, with the mixing weights, still perform slightly better than the proposed exponentiated models.

Table 3. Comparisons of different models, with the Danish fire insurance data.

Model	k [1]	NLL	AIC	BIC	KS	$VaR_{0.90}$	$VaR_{0.95}$
Empirical Estimate						5.08	8.41
IG-Pareto (One-Parameter)	1	6983.8	13,969.6	13,975.5	0.575	5.24×10^5	3.59×10^7
Exponentiated IG-Pareto	2	4287.7	8591.0	8590.0	0.136	13.82	32.48
exp-Pareto (One-Parameter)	1	5878.0	11,758.0	11,763.8	0.333	325.62	2359.69
Exponentiated exp-Pareto	2	3961.0	7926.0	7937.7	0.057	4.42	6.89
Weibull-Inverse Weibull [2]	4	3820.0	7648.0	7671.3	0.021	5.08	8.02
Weibull	2	5270.5	10,545.0	10,556.6	0.409	4.57	5.98
IG	2	4097.9	8199.8	8211.4	0.358	4.67	6.41

[1] k is the number of parameters in the model. [2] The composite model has an additional weight parameter ϕ [5].

5.2. Norwegian Fire Insurance Data for 1990

The Norwegian fire insurance data was utilized in the previous literature, to check the performance of different loss models [4,8,19]. The dataset contains 9181 claims from a Norwegian insurance company. The unit of the claims is in 1000s of Norwegian Krones (NOK). We were able to access the data via R package *ReIns* [20]. Since we do not know if the data was inflation adjusted, we chose only the claims from 1990, for the analysis.

Figure 6 is the histogram of the data set. There are 628 reported claims from the 1990 in the dataset. For analysis, we scaled the data, so that the claim values had a unit of millions of Norwegian Krones (NOK). The detailed summary of the dataset is provided in Table 4.

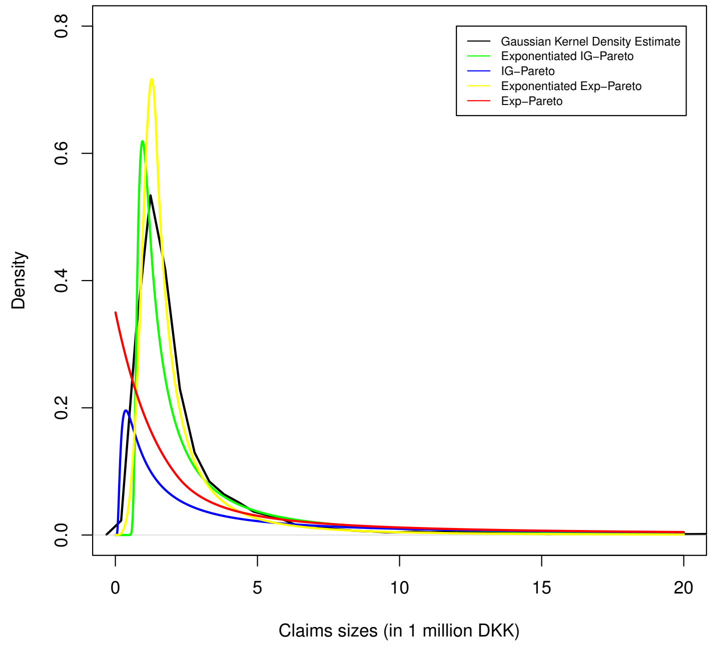

Figure 4. Kernel density estimate of Danish fire insurance data, with corresponding exponentiated exp-Pareto, exp-Pareto, exponentiated IG-Pareto, and IG-Pareto fit.

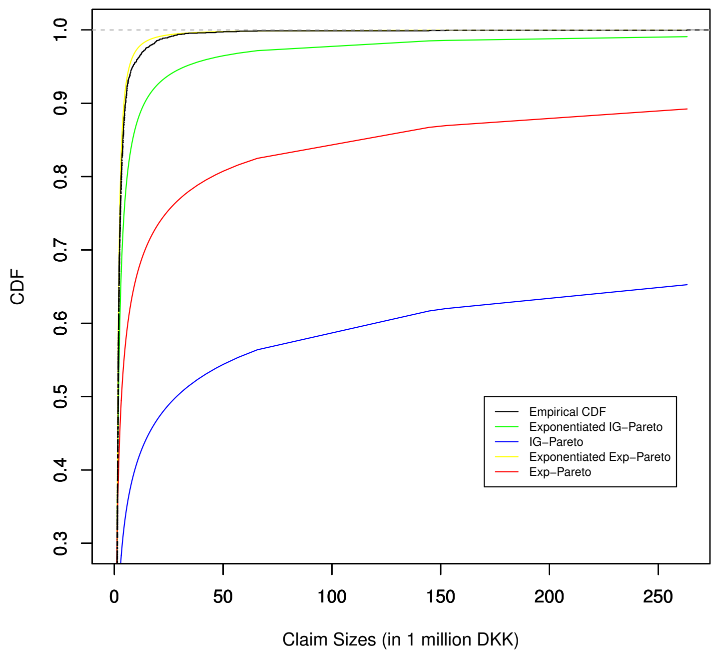

Figure 5. The empirical CDF of Danish fire insurance data and the fitted CDF, of corresponding exponentiated IG-Pareto, IG-Pareto, exponentiated exp-Pareto, and exp-Pareto model fit.

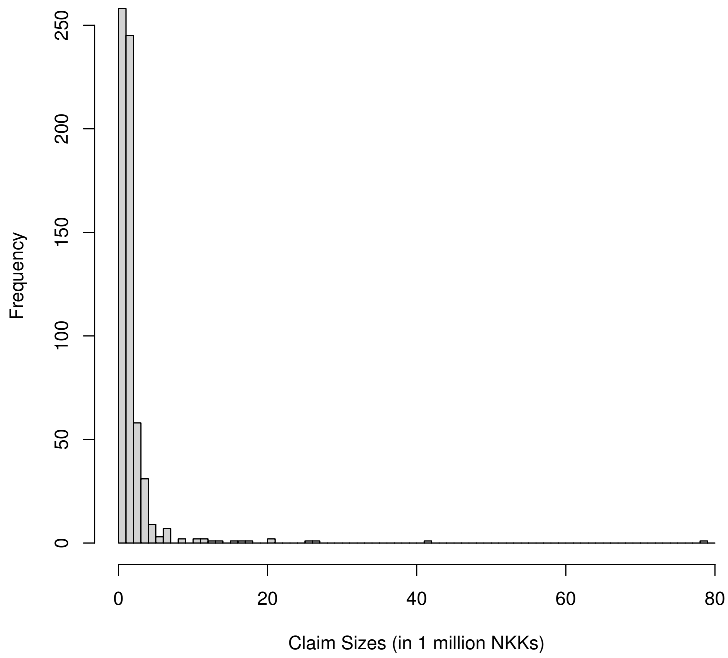

Figure 6. Histogram of Norwegian fire insurance data for 1990.

Table 4. Summary statistics for Norwegian fire insurance claims, during 1990 (in 1 million NOKs).

Sample Size	Mean	SD	Minimum	Q1	Q2	Q3	Maximum
628	1.97	4.26	0.50	0.79	1.15	1.81	78.54

Table 5 summarizes the comparisons of different models, when fitting to the Norwegian fire insurance data. Like the results that were obtained for the previous dataset, the two proposed exponentiated-composite models significantly overperform the original one-parameter composite models, with regard to all measures. This is, also, illustrated visually in Figure 7. The two exponentiated models provided closer fits the Gaussian kernel density estimate of the dataset, compared to the original one-parameter composite models. Both of the two proposed two-parameter exponentiated models perform better than the Weibull and IG models. Figure 8 shows the comparison of the fitted CDF of the exponentiated-composite models, the original composite models, and the empirical CDF of the data. Both of the exponentiated-composite models demonstrated a good fit to the empirical CDF, while both of the original composite models cannot fit the data properly. In addition, among all seven models that we chose for the real data application, the exponentiated inverse Gamma-Pareto model demonstrates the best performance, in terms of BIC.

5.3. Society of Actuaries (SOA) Group Medical Insurance Claims Data: Year 1991

The Society of Actuaries (SOA) group medical insurance claims is a dataset that contains 171,100 claims that exceed USD 25,000 from year 1991 to 1992. There is no information for the inflation adjustment for this dataset, so we choose the 75,789 claims from year 1991 in our analysis. Figure 9 is the histogram of the dataset. The detailed summary of the dataset is presented in Table 6. The unit of the claims was scaled to USD 10,000, for analysis purposes.

The performances of different models, when fitting to the SOA group medical claims data, are presented in Table 7. The two proposed exponentiated-composite models, still, demonstrate significantly better performances, compared to the original one-parameter composite models, with respect to all GoF measures. Figure 10, also, confirms that the two exponentiated-composite models provided better fit to the kernel-density estimate of the data, in comparison with the two original composite models. Both of the two proposed two-parameter exponentiated-composite models perform better than the two-parameter Weibull and IG models. The comparison of the fitted CDF of the exponentiated-composite models, original composite models, and the empirical CDF is presented in Figure 11. Again, both of the original composite models cannot provide a proper fit to the data, while the exponentiated-composite models demonstrated very close fits to the empirical CDF. Within all the chosen models, the exponentiated IG-Pareto model performed the best, concerning all the GoF measures. In addition, the VaR estimates for the fitted exponentiated IG-Pareto model, at the level of 0.90 and 0.95, are very close to the empirical VaR values. This suggests that the model is capable of capturing the data features at the extreme upper tail.

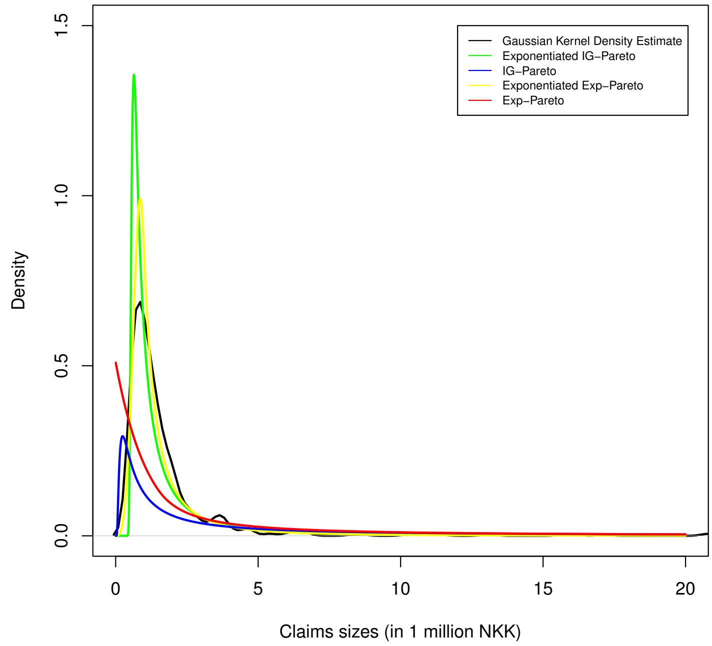

Figure 7. Kernel density estimate of Norwegian fire insurance data, for 1990, with corresponding exponentiated exp-Pareto, exp-Pareto, exponentiated IG-Pareto, and IG-Pareto fit.

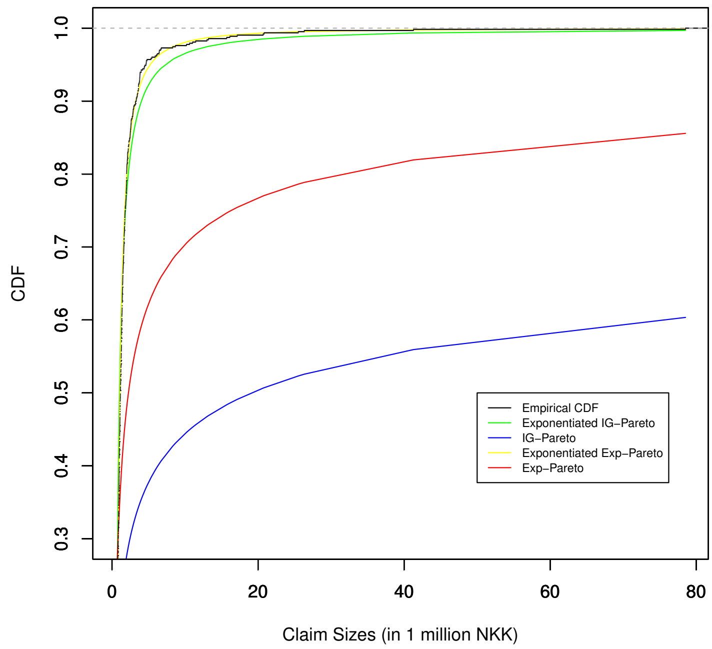

Figure 8. The empirical CDF of Norwegian fire insurance data and the fitted CDF of corresponding exponentiated IG-Pareto, IG-Pareto, exponentiated exp-Pareto, and exp-Pareto model fit.

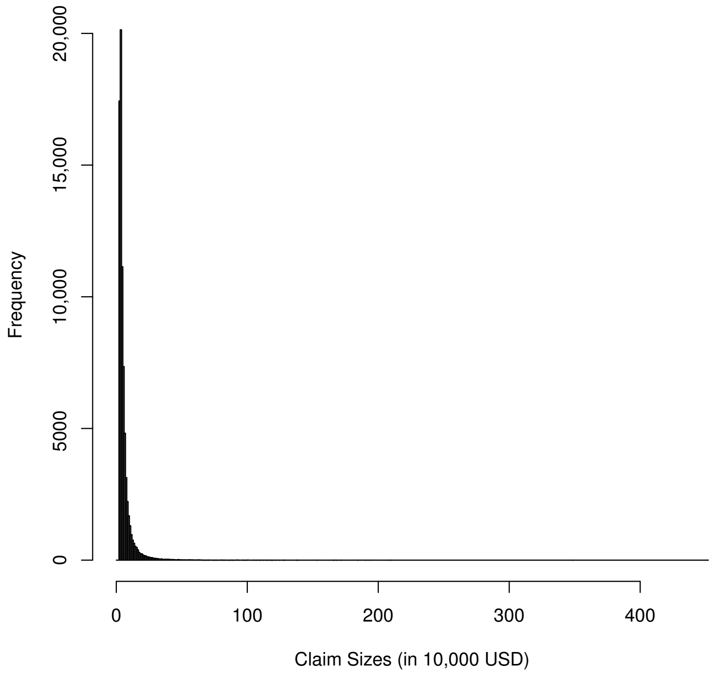

Figure 9. Histogram of SOA group medical insurance claims data, for 1991.

Table 5. Comparisons of different models with the Norwegian fire insurance data, for 1990.

Model	k [1]	NLL	AIC	BIC	KS	$VaR_{0.90}$	$VaR_{0.95}$
Empirical Estimate						3.25	4.59
IG-Pareto (One-Parameter)	1	1503.7	3009.4	2013.9	0.591	3.50×10^5	2.40×10^7
Exponentiated IG-Pareto	2	755.6	1515.1	1524.0	0.097	4.02	7.28
exp-Pareto (One-Parameter)	1	1225.6	2453.3	2457.7	0.355	223.64	1620.64
Exponentiated exp-Pareto	2	772.2	1548.4	1557.3	0.053	3.26	5.23
Weibull-Inverse Weibull [2]	4	751.0	1509.9	1527.7	0.038	3.02	4.36
Weibull	2	1054.4	2112.8	2121.7	0.232	4.57	5.98
IG	2	773.9	1551.7	1560.6	0.049	3.09	4.24

[1] k is the number of parameters in the model. [2] The composite model has an additional weight parameter ϕ [5].

Table 6. Summary statistics for SOA group medical insurance claims, for 1991 (in 10,000 USD).

Sample Size	Mean	SD	Minimum	Q1	Q2	Q3	Maximum
75,789	5.84	6.60	2.50	3.05	4.02	6.13	451.84

Table 7. Comparisons of different models, with the SOA group medical claims insurance data, for 1991.

Model	k [1]	NLL	AIC	BIC	KS	$VaR_{0.90}$	$VaR_{0.95}$
Empirical Estimate						10.18	14.76
IG-Pareto (One-Parameter)	1	277,440.9	347,243.2	347,261.7	0.602	1.36×10^6	9.36×10^7
Exponentiated IG-Pareto	2	160,836.6	321,677.2	321,695.7	0.035	9.77	14.54
exp-Pareto (One-Parameter)	1	242,318.1	484,638.2	484,647.4	0.369	835.46	6054.35
Exponentiated exp-Pareto	2	168,683.5	337,371.0	337,389.5	0.085	8.69	12.29
Weibull-Inverse Weibull [2]	4	167,745.8	335,499.6	335,536.5	0.070	8.67	11.40
Weibull	2	204,223.9	408,451.8	408,470.3	0.256	12.24	15.01
IG	2	173,619.6	554,883.8	554,893.0	0.099	9.29	11.78

[1] k is the number of parameters in the model. [2] The composite model has an additional weight parameter ϕ [5].

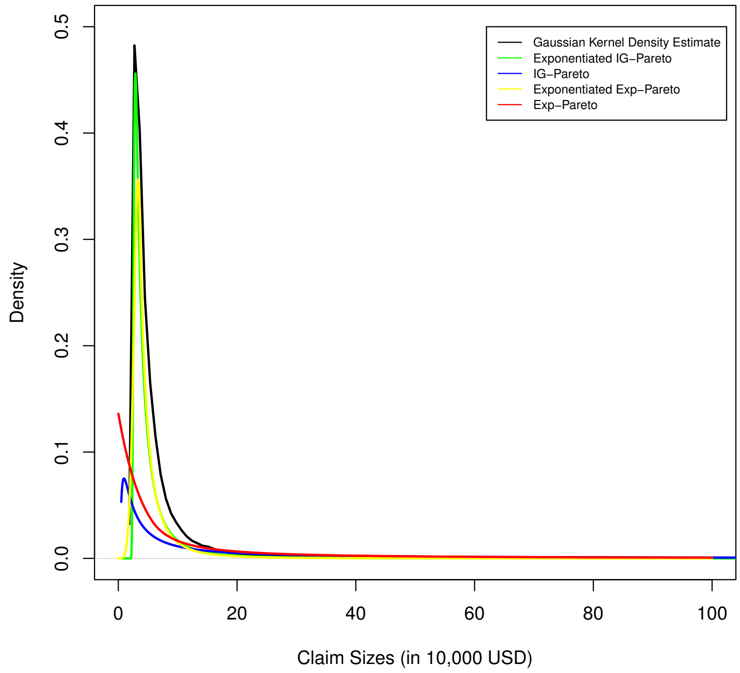

Figure 10. Kernel density estimate of SOA group medical claims data, with corresponding exponentiated exp-Pareto, exp-Pareto, exponentiated IG-Pareto, and IG-Pareto fit.

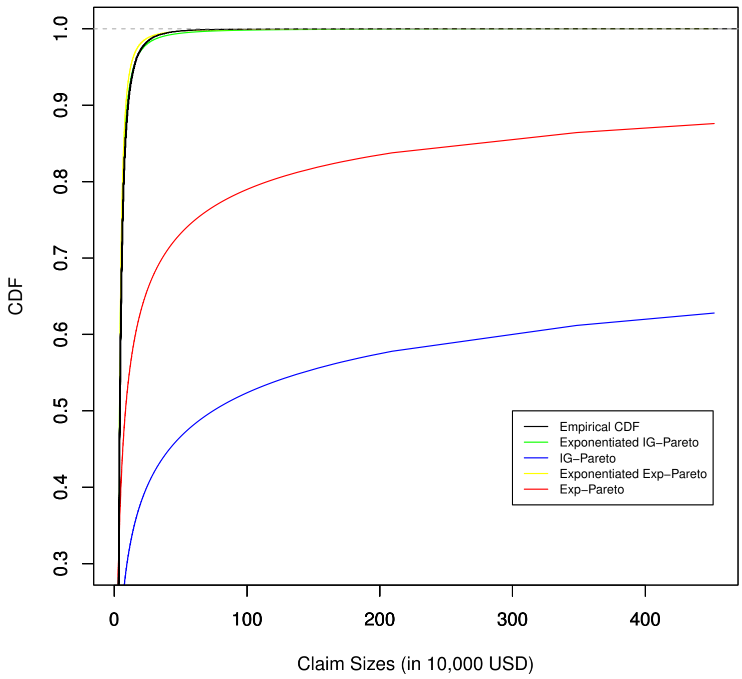

Figure 11. The empirical CDF of SOA group medical claims data and the fitted CDF of corresponding exponentiated IG-Pareto, IG-Pareto, exponentiated exp-Pareto, and exp-Pareto model fit.

5.4. Estimated Parameter Values for the Three Data Sets

For the three datasets, considered in the three previous subsections, we provided estimated parameter values for the exponentiated IG-Pareto model and exponentiated exp-Pareto model, in Table 8. Furthermore, their estimated standard errors are presented in parenthesis next to the corresponding estimates.

Table 8. The parameter estimates and estimated standard error of the exponentiated-composite models, when fitted to the three datasets.

Dataset	Estimated Parameters	Exponentiated IG-Pareto (se)	Exponentiated Exp-Pareto (se)
Danish Fire Insurance Data	$\hat{\theta}$	2.81 (0.15)	5.22 (0.24)
	$\hat{\eta}$	4.95 (0.07)	4.47 (0.09)
Norwegian Fire Insurance Data	$\hat{\theta}$	0.13 (0.02)	0.96 (0.09)
	$\hat{\eta}$	7.13 (0.27)	4.19 (0.16)
SOA Group Medical Claims Data	$\hat{\theta}$	216,610 (8131.23)	1596.72 (38.98)
	$\hat{\eta}$	10.64 (0.04)	5.72 (0.02)

6. Discussion and Concluding Remarks

In this paper, we proposed a generalized family of exponentiated-composite distributions. The motivation of proposing such a family is to improve the flexibility of the original composite distribution, by introducing an exponent parameter. Similar to how the Weibull distribution was developed, based on exponential distribution, we introduce the exponent parameter, by exponentiating a general random variable associated with a composite distribution, defined in Section 2.1. We proved that an exponentiated-composite distribution is, still, a composite distribution and we derived some mathematical properties of this new family of distribution, including the raw moments and the limited moments. The two-parameter exponentiated IG-Pareto and exponentiated exp-Pareto model were discussed, as the special models within this family. We, also, provided a method to find the estimates of the parameters in an exponentiated-composite model. The simulations in Section 4 showed that the method has the ability to identify the true parameters, for an exponentiated exp-Pareto model. We assess the performances of the two-parameter exponentiated IG-Pareto and exponentiated exp-Pareto model, with the three insurance datasets, and compared their performances to some existing models, in the past literature. For the Norwegian fire insurance data, for 1990, the exponentiated IG-Pareto model demonstrates the best performance among all he models we chose, in terms of BIC. When fitting to the SOA group medical claims data, the exponentiated IG-Pareto model outperformed all the other models we chose for comparison. We foresee that this family of exponentiated models will significantly increase the abundance of the composite models. Since we, also, proved the exponentiating composite models are, still, composite models, this family can be extended to the exponentiated-composite models, with the mixing weights.

Author Contributions: Conceptualization, B.L. and M.M.A.A.; methodology, B.L. and M.M.A.A.; software, B.L.; validation, B.L. and M.M.A.A.; formal analysis, B.L.; investigation, B.L. and M.M.A.A.; resources, B.L.; writing—original draft preparation, B.L.; writing—review and editing, B.L. and M.M.A.A.; supervision, M.M.A.A.; project administration, B.L. and M.M.A.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: All data related to this study are available in public, via R package 'ReIns' and 'SMPracticals'.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

pdf	probability density function
IG	inverse Gamma
exp-Pareto	exponential-Pareto
i.i.d.	independent and identically distributed
SE	standard error
NLL	negative log-likelihood
AIC	Akaike information criterion

BIC	Bayesian information criterion
MLE	maximum likelihood estimate
CDF	cumulative distribution function
GoF	goodness-of-fit
VaR	value-at-risk
DKK	Danish Krones
NOK	Norwegian Krones

References

1. Cooray, K.; Ananda, M.M.A. Modeling actuarial data with a composite lognormal-Pareto model. *Scand. Actuar. J.* **2005**, *2005*, 321–334. [CrossRef]
2. Scollnik, D.P.M. On composite lognormal-Pareto models. *Scand. Actuar. J.* **2007**, *2007*, 20–33. [CrossRef]
3. Abu Bakar, S.; Hamzah, N.; Maghsoudi, M.; Nadarajah, S. Modeling loss data using composite models. *Insur. Math. Econ.* **2015**, *61*, 146–154. [CrossRef]
4. Brazauskas, V.; Kleefeld, A. Modeling Severity and Measuring Tail Risk of Norwegian Fire Claims. *N. Am. Actuar. J.* **2016**, *20*, 1–16. [CrossRef]
5. Grün, B.; Miljkovic, T. Extending composite loss models using a general framework of advanced computational tools. *Scand. Actuar. J.* **2019**, *2019*, 642–660. [CrossRef]
6. Aminzadeh, M.S.; Deng, M. Bayesian predictive modeling for Inverse Gamma-Pareto composite distribution. *Commun. Stat. Theory Methods* **2019**, *48*, 1938–1954. [CrossRef]
7. Deng, M.; Aminzadeh, M.S. Bayesian predictive analysis for Weibull-Pareto composite model with an application to insurance data. *Commun. Stat. Simul. Comput.* **2019**, *51*, 1–27. [CrossRef]
8. Liu, B.; Ananda, M.M.A. Analyzing insurance data with an exponentiated composite Inverse-Gamma Pareto Model. *Commun. Stat. Theory Methods* **2022**. [CrossRef]
9. Mutali, S.; Vernic, R. On the composite Lognormal–Pareto distribution with uncertain threshold. *Commun. Stat. Simul. Comput.* **2020**, 1–17. [CrossRef]
10. Cooray, K.; Gunasekera, S.; Ananda, M. Weibull and inverse Weibull composite distribution for modeling reliability data. *MAS* **2010**, *5*, 109–115. [CrossRef]
11. Kim, D.I.; Kwon, H.H.; Han, D. Bias correction of daily precipitation over South Korea from the long-term reanalysis using a composite Gamma-Pareto distribution approach. *Hydrol. Res.* **2019**, *50*, 1138–1161. [CrossRef]
12. Cooray, K. The Weibull–Pareto Composite Family with Applications to the Analysis of Unimodal Failure Rate Data. *Commun. Stat. Theory Methods* **2009**, *38*, 1901–1915. [CrossRef]
13. Teodorescu, S.; Vernic, R. A Composite Exponential-Pareto Distribution. *Analele Ştiinţifice ale Universităţii "Ovidius" Constanţa Seria Matematică* **2006**, *14*, 99–108.
14. Bahnemann, D. Distributions for Actuaries. In *CAS Monograph Series No. 2*; Casualty Actuarial Society: Arlington, TX, USA, 2015.
15. Sablica, L.; Hornik, K. mistr: A Computational Framework for Mixture and Composite Distributions. *R J.* **2020**, *12*, 283–299. [CrossRef]
16. Burnham, K.P.; Anderson, D.R. *Model Selection and Multimodel Inference*, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 2002.
17. Smirnov, N. Table for Estimating the Goodness of Fit of Empirical Distributions. *Ann. Math. Stat.* **1948**, *19*, 279–281. [CrossRef]
18. Davison, A. SMPracticals: Practicals for Use with Davison (2003) Statistical Models, R Package Version 1.4-3. 2019. Available online: https://cran.r-project.org/web/packages/SMPracticals/ (accessed on 31 March 2022).
19. Beirlant, J.; Matthys, G.; Dierckx, G. Heavy-Tailed Distributions and Rating. *ASTIN Bull. J. IAA* **2001**, *31*, 37–58. [CrossRef]
20. Reynkens, T.; Verbelen, R. ReIns: Functions from "Reinsurance: Actuarial and Statistical Aspects", R Package Version 1.0.10. 2020. Available online: https://cran.r-project.org/web/packages/ReIns/ (accessed on 31 March 2022).