Hindawi Publishing Corporation
Journal of Probability and Statistics
Volume 2009, Article ID 364901, 10 pages
doi:10.1155/2009/364901

Research Article

From the General Affine Transform Family to a Pareto Type IV Model

Werner Hürlimann

Feldstrasse 145, 8004 Zürich, Switzerland

Correspondence should be addressed to Werner Hürlimann, whurlimann@bluewin.ch

Received 30 March 2009; Revised 30 September 2009; Accepted 15 October 2009

Recommended by José María Sarabia

The analytical form of general affine transform families with given maximum likelihood estimators for the affine parameters is determined. In this context, the simultaneous maximum likelihood equations of the affine parameters in the generalised Pareto distribution cannot have a common solution. This pathological situation is removed by extending it to a four parameter family, called Pareto type IV model.

1. Introduction

Based on [1], the author has studied the general affine transform X of the random variable Y defined by $X = U[A(\alpha)+B(\alpha)\cdot\psi(Y)]$, where $\psi(x)$ and $U(x)$ are twice differentiable monotone increasing functions, and $A(\alpha)$, $B(\alpha)$ are deterministic functions of the affine parameter vector α such that $B(\alpha) > 0$. The work in [2] determines exact maximum likelihood estimators of parameters in order statistics distributions with exponential, Pareto, and Weibull parent distributions. The article [3] recovers the older result by the work in [4] that the Pareto is an exponential transform, and also notes that the latter result is not restricted to the Pareto, but applies to a lot of distributions like the truncated Cauchy, Gompertz, log-logistic, para-logistic, inverse Weibull, and log-Laplace.

A further contribution in this area is offered. Based on the method introduced in [5], we determine the analytical form that parametric models may take for specific maximum likelihood estimators of the affine parameters in a general affine transform family. Applied to the generalised Pareto distribution, of great importance in extreme value theory and its applications (e.g., [6, 7]), one observes that the simultaneous maximum likelihood equations of the affine parameters cannot have a common solution. Therefore, the highly desirable maximum likelihood method is not applicable to this distribution. Fortunately, this pathological situation can be removed by enlarging the generalised Pareto to a four-parameter family. The resulting new family, called Pareto type IV model, includes as special cases the generalised Pareto and the Beta of type II. Finally, it is worthwhile to mention the construction of alternative statistical models of Pareto type II and III in [8], and of type IV in [9]. A recent discussion of the Pareto type III is [10] and a useful monograph including Pareto type distributions is [11]. This paper is organized as follows.

Section 2 recalls the general affine transform family (GATF) and its relevance. Our main result concerns the possible form GATF models may take given specific maximum likelihood estimators (MLE) for their affine parameters and is derived in Section 3. Section 4 shows that our method does not apply to the generalised Pareto distribution and introduces the new Pareto type IV model. Section 5 concludes and gives a short outlook on further research.

2. General Affine Transform Families

Let X, Y be random variables with distribution functions F_X, F_Y and densities f_X, f_Y (provided they exist). Suppose that the distributions and densities depend on a parameter vector $\theta = (\alpha, \gamma)$ with values in the parameter space $\Theta \subset R^m$, where $\alpha = (\alpha_1, \ldots, \alpha_r)$ is a vector of *affine parameters*, $\gamma = (\gamma_1, \ldots, \gamma_s)$ is a vector of *shape parameters*, and $m = r + s$. We assume that the functions $\psi(x)$ and $U(x)$ are continuous twice-differentiable monotone increasing with inverses $\varphi(x) = \psi^{-1}(x)$ and $T(x) = U^{-1}(x)$. Moreover, these functions do not depend on α but may depend on γ.

Definition 2.1. The *general affine transform* X of Y (GATF) is the random variable defined by $X = U[A(\alpha)+B(\alpha)\cdot\psi(Y)]$ via a three-stage transformation. First, Y is nonlinearly transformed to $\psi(Y)$, then positively linear transformed to $T(Y) = A(\alpha) + B(\alpha) \cdot \psi(Y)$, with $B(\alpha) > 0$, and again nonlinearly transformed to $X = U[T(Y)]$. The constants $A(\alpha)$ and $B(\alpha)$ are called *location* and *scale* parameters. A *GATF family* $F\{Y\} = \{X = U[A(\alpha) + B(\alpha) \cdot \psi(Y)] \sim F_X(x;\theta), \theta = (\alpha, \gamma) \in \Theta\}$ is a set of parameterised GATF X of Y whose distributions and densities satisfy the relationships

$$F_X(x) = F_Y\left\{\varphi\left[\frac{T(x) - A(\alpha)}{B(\alpha)}\right]\right\}, \tag{2.1}$$

$$f_X(x) = \frac{1}{B(\alpha)} \cdot T'(x) \cdot \varphi'\left[\frac{T(x) - A(\alpha)}{B(\alpha)}\right] \cdot f_Y\left\{\varphi\left[\frac{T(x) - A(\alpha)}{B(\alpha)}\right]\right\}. \tag{2.2}$$

In applications, very often special cases are most useful. Using [1, Table 1], the main types are summarized in [3, Table 2.1]. Some typical examples illustrate the relevance of the GATF as the generalised Pareto and the gxh-family [3, Examples 2.1 and 2.2].

3. GATF Families with Prescribed Maximum Likelihood Estimators

Consider a random sample $\xi = (X_1, \ldots, X_n)$ of size n, where X_i are independent and identically distributed random variables, and denote the common random variable by X. For a real function $H(x)$, we define and denote the *mean value* of $H(\xi)$ by

$$\overline{H(\xi)} = \frac{1}{n}\sum_{i=1}^{n} H(X_i). \tag{3.1}$$

It is assumed that sample mean value equations like $\overline{H(\xi/\widehat{\alpha})} = 1$ have a unique solution $\widehat{\alpha} = \widehat{\alpha}(\xi, H)$. Our main result characterizes GATF families by the form of the maximum likelihood estimators for their affine parameters. The proof makes use in [12, Theorem 2.2].

Theorem 3.1. *Given is a GATF $X = U[A(\alpha) + B(\alpha) \cdot \psi(Y)]$ with support $I_X = [a_X, b_X]$ and affine parameter vector $\alpha = (\alpha_1, \ldots, \alpha_r) \in \Theta \subset R^r$. Suppose that the distribution function $F_X(x)$ of X is twice differentiable, and that the MLE of the kth affine parameter α_k is solution of one of the following mean value equations.*

Case 1.

$$B_k = \frac{\partial B(\alpha)}{\partial \alpha_k} \neq 0, \qquad A_k = \frac{\partial A(\alpha)}{\partial \alpha_k} \text{ arbitrary}, \quad k \in \{1, \ldots, r_1\},$$
$$\overline{S_k\left(\frac{T(\xi) - A(\widehat{\alpha})}{B(\widehat{\alpha})} + \frac{A_k}{B_k}\right)} = 1, \tag{3.2}$$

with some real function $S_k(x)$.

Case 2.

$$B_k = \frac{\partial B(\alpha)}{\partial \alpha_k} \equiv 0, \qquad A_k = \frac{\partial A(\alpha)}{\partial \alpha_k} \neq 0, \quad k \in \{r_1 + 1, \ldots, r\},$$
$$\overline{L_k\left(\frac{T(\xi) - A(\widehat{\alpha})}{B(\widehat{\alpha})}\right)} = 0, \tag{3.3}$$

with some real function $L_k(x)$.

Then there exists a twice-differentiable and monotone increasing function $Q_k(x)$ with derivative $q_k(x) = Q_k'(x)$, and constants $c_k, d_k \neq 0$ such that

$$c_k S_k(x) + 1 - c_k = -x \cdot \frac{d}{dx} \ln\{q_k(x)\}, \quad \textit{in Case } 1, \tag{3.4}$$

$$d_k L_k(x) = -\frac{d}{dx} \ln\{q_k(x)\}, \quad \textit{in Case } 2. \tag{3.5}$$

Furthermore, for simultaneous maximum likelihood estimation of the affine parameters, the following compatibility conditions must be satisfied:

$$\left(x + \frac{A_j}{B_j}\right) \cdot \left(c_i S_i\left(x + \frac{A_i}{B_i}\right) + 1 - c_i\right) = \left(x + \frac{A_i}{B_i}\right) \cdot \left(c_j S_j\left(x + \frac{A_j}{B_j}\right) + 1 - c_j\right), \quad i, j \in \{1, \ldots, r_1\}, \tag{3.6}$$

$$c_i S_i\left(x + \frac{A_i}{B_i}\right) + 1 - c_i = \left(x + \frac{A_i}{B_i}\right) \cdot d_j L_j(x), \quad i \in \{1, \ldots, r_1\},\ j \in \{r_1 + 1, \ldots, r\}, \tag{3.7}$$

$$d_i L_i(x) = d_j L_j(x), \quad i, j \in \{r_1 + 1, \ldots, r\}. \tag{3.8}$$

Under these conditions, the distribution function has the unique representation

$$F_X(x) = \frac{Q_i((T(x)-A)/B + (A_i/B_i)) - Q_i((T(a_X)-A)/B + (A_i/B_i))}{Q_i((T(b_X)-A)/B + (A_i/B_i)) - Q_i((T(a_X)-A)/B + (A_i/B_i))}$$
$$= \frac{Q_j((T(x)-A)/B) - Q_j(T(a_X)-A/B)}{Q_j(T(b_X)-A/B) - Q_j(T(a_X)-A/B)}, \tag{3.9}$$

for all $x \in I_X = [a_X, b_X]$, $i \in \{1,\dots,r_1\}$, $j \in \{r_1+1,\dots,r\}$.

Proof. We proceed as in [5, proof of Theorem 2.1].

Case 1 ($k \in \{1,\dots,r_1\}$). Using (2.2) and the relations $Y = \varphi((T(X)-A)/B), \varphi'[\psi(Y)] = \psi'(Y)^{-1}$, one obtains for the negative of the random log-likelihood of X the expression

$$-\ell(X) = \ln B(\alpha) - \ln T'(X) + \ln \psi'(Y) - \ln f_Y(Y). \tag{3.10}$$

Denoting partial derivatives with respect to α_k with a lower index k and making use of

$$Y_k = \varphi'\left(\frac{T(X)-A}{B}\right) \cdot \frac{-A_k B - (T(X)-A)B_k}{B^2} = -\frac{A_k + B_k\psi(Y)}{B\psi'(Y)}, \tag{3.11}$$

one obtains from (3.10) the expression for the partial derivative

$$-\frac{B}{B_k} \cdot \ell_k(X) = 1 - \left(\frac{\psi(Y) + (A_k/B_k)}{\psi'(Y)}\right) \cdot \left(\frac{\psi''(Y)}{\psi'(Y)} - \frac{d}{dY}\ln\{f_Y(Y)\}\right). \tag{3.12}$$

By assumption (3.2), one has using [12, Theorem 2.2] that

$$-\frac{B}{B_k} \cdot \ell_k(X) = c_k \cdot \left\{1 - S_k\left(\psi(Y) + \frac{A_k}{B_k}\right)\right\} \tag{3.13}$$

for some constant $c_k \neq 0$. By comparison $y(x) = \psi(x) + (A_k/B_k)$ solves the second-order differential equation

$$\frac{y''}{y'} - \{c_k S_k(y) + 1 - c_k\} \cdot \frac{y'}{y} = \frac{d}{dx}\ln\{f_Y(x)\}. \tag{3.14}$$

Setting $g_k(x) = (c_k S_k(x) + (1-c_k))/x$ and multiplying with y', this simplifies to

$$y'' - \frac{d}{dx}\ln\{f_Y(x)\} \cdot y' - g_k(y) \cdot y'^2 = 0. \tag{3.15}$$

Transform it to the equivalent system of first-order equations in $(y_1 = y, y_2)$ [13, Chapter 19]:

$$y_1' = y_2, \qquad y_2' = \frac{d}{dx}\ln\{f_Y(x)\} \cdot y_2 + g_k(y_1) \cdot y_2^2. \tag{3.16}$$

The second differential equation is of Bernoulli type [13, Chapter 2]. Setting $y_2 = z_2^{-1}$, this is equivalent to the simpler system in (y_1, z_2):

$$y_1' = z_2^{-1}, \qquad z_2' = -\frac{d}{dx}\ln\{f_Y(x)\} \cdot z_2 + g_k(y_1). \tag{3.17}$$

The second equation is linear inhomogeneous of first order and has the homogeneous solution $z_2 = C_k \cdot f_Y(x)^{-1}$. By variation of the constant, one sees that $C_k'(x) = -g_k(y_1) \cdot f_Y(x)$. On the other side, from the first equation in (3.17), one has $y' = y_1' = z_2^{-1} = C_k(x)^{-1} \cdot f_Y(x)$, hence $f_Y(x) = y_1' \cdot C_k(x)$. Together, this shows the following separated differential equation:

$$\frac{d}{dx}\ln\{C_k(x)\} = -g_k(y) \cdot y'. \tag{3.18}$$

Assume momentary that $g_k(x)$ has an integral such that $G_k'(x) = g_k(x)$ for some $G_k(x)$. Then, $(d/dx)\ln\{C_k(x)\} = -(d/dx)G_k(y)$ has the solution $C_k(x) = C_k^{-1} \cdot \exp\{-G(y)\}$, $C_k > 0$. It follows that the general solution of the second differential equation in (3.17) is given by

$$z_2 = \frac{\exp\{-G_k(y)\}}{C_k f_Y(x)}. \tag{3.19}$$

The first differential equation in (3.17) implies the separated differential equation

$$y' \cdot \exp\{-G_k(y)\} = C_k \cdot f_Y(x). \tag{3.20}$$

Assume momentary that there exists a twice-differentiable function $Q_k(x)$ such that $G_k(x) = -\ln\{Q_k'(x)\}(g_k(x) = G_k'(x) = -(Q_k''(x)/Q_k'(x)))$. The general solution to (3.20) yields the relationship

$$F_Y(x) = \frac{1}{C_k}\{Q_k(y) + D_k\}, \quad C_k > 0,\ D_k \in R. \tag{3.21}$$

Setting $x = Y$ and using that $y(x) = \varphi(Y) + (A_k/B_k) = (T(X) - A)/B + (A_k/B_k)$, one gets the random relation $F_Y(Y) = (1/C_k)\{Q_k((T(X) - A)/B + (A_k/B_k)) + D_k\}$, which implies by (2.1) that

$$F_X(x) = \frac{1}{C_k}\left\{Q_k\left(\frac{T(x) - A}{B} + \frac{A_k}{B_k}\right) + D_k\right\}, \quad x \in I_X. \tag{3.22}$$

Setting $q_k(x) = Q_k'(x)$, one obtains the density function

$$f_X(x) = \frac{T'(x)}{BC_k} q_k\left(\frac{T(x) - A}{B} + \frac{A_k}{B_k}\right), \quad x \in I_X. \tag{3.23}$$

The side conditions $\int_{a_X}^{b_X} f_X(x)dx = 1$, $F_X(b_X) = 1$, imply that the constants are determined by

$$C_k = Q_k\left(\frac{T(b_X) - A}{B} + \frac{A_k}{B_k}\right) - Q_k\left(\frac{T(a_X) - A}{B} + \frac{A_k}{B_k}\right), \qquad D_k = -Q_k\left(\frac{T(a_X) - A}{B} + \frac{A_k}{B_k}\right). \tag{3.24}$$

The validity of the representation (3.9) for $i \in \{1, \ldots, r_1\}$ is shown. Since $F_Y(x)$ has been assumed twice differentiable, so is $Q_k(x)$, and

$$c_k S_k(x) + 1 - c_k = x g_k(x) = x G_k'(x) = -x \cdot \frac{d}{dx} \ln\{q_k(x)\}, \tag{3.25}$$

as claimed in (3.4). In particular, the two momentary assumptions made above, that is, $g_k(x) = G_k'(x)$ and $G_k(x) = -\ln\{Q_k'(x)\}$, are fulfilled.

Case 2 ($k \in \{r_1 + 1, \ldots, r\}$). Since $B_k \equiv 0$, one has similarly to (3.11) the relationship

$$Y_k = -\frac{A_k}{B\psi'(Y)}. \tag{3.26}$$

From (3.10), one obtains for the partial derivative of the random log-likelihood the relation

$$-\frac{B}{A_k} \cdot \ell_k(X) = \frac{1}{\psi'(Y)} \cdot \left(\frac{\psi''(Y)}{\psi'(Y)} - \frac{d}{dY} \ln\{f_Y(Y)\}\right). \tag{3.27}$$

By assumption (3.2) and again in [12, Theorem 2.2], one has

$$-\frac{B}{A_k} \cdot \ell_k(X) = d_k \cdot L_k(\psi(Y)) \tag{3.28}$$

for some constant $d_k \neq 0$. Through comparison, it follows that $y(x) = \psi(x)$ must solve

$$y'' - \frac{d}{dx} \ln\{f_Y(x)\} \cdot y' - d_k \cdot L_k(y) \cdot y'^2 = 0. \tag{3.29}$$

Proceeding as in Case 1, one obtains a twice-differentiable function $Q_k(x)$, with derivative $q_k(x) = Q_k'(x)$, such that $d_k L_k(x) = -(d/dx)\ln\{q_k(x)\}$ and $F_Y(x) = (1/C_k)\{Q_k(y) + D_k\}, C_k > 0, D_k \in R$. As in Case 1, one concludes that (3.9) for $j \in \{r_1 + 1, \ldots, r\}$ must hold.

It remains to show the compatibility conditions (3.6)–(3.8). Through differentiation of (3.9), one obtains the probability density functions

$$f_X(x) = \frac{T'(x)}{BC_i} q_i\left(\frac{T(x) - A}{B} + \frac{A_i}{B_i}\right) = \frac{T'(x)}{BC_j} q_j\left(\frac{T(x) - A}{B}\right), \tag{3.30}$$

for all $x \in I_X$, $\quad i \in \{1, \ldots, r_1\}$, $\quad j \in \{r_1 + 1, \ldots, r\}$. Three subcases are possible.

Subcase 1 ($i, j \in \{1, \ldots, r_1\}$). From (3.30), one gets that $q_j(x + (A_i/B_i)) = C \cdot q_i(x + (A_j/B_j))$ with $C = C_j/C_i$. Using (3.4), one obtains without difficulty the compatibility condition (3.6).

Subcase 2 ($i \in \{1, \ldots, r_1\}, j \in \{r_1+1, \ldots, r\}$). From (3.30), one sees that $q_j(x) = C \cdot q_i(x + (A_j/B_j))$ with $C = (C_j/C_i)$. Using (3.4) and (3.5), one shows without difficulty condition (3.7).

Subcase 3 ($i, j \in \{r_1 + 1, \ldots, r\}$). From (3.30), one obtains that $q_j(x) = C \cdot q_i(x)$ with $C = C_j/C_i$. Using (3.5), one shows without difficulty condition (3.8). The proof of Theorem 3.1 is complete. □

4. A Pareto Type IV Model

The generalised Pareto distribution is the GATF defined by $X = A(\alpha) + B(\alpha) \cdot \varphi(Y)$ with $\varphi(x) = \exp(\gamma_1 x), \gamma_1 > 0$, Y exponential with mean one, $A(\alpha) = \alpha_2 - \alpha_1$, $B(\alpha) = \alpha_1$, $\alpha = (\alpha_1, \alpha_2) \in R_+^2$, $\theta = (\alpha_1, \alpha_2, \gamma_1) \in \Theta = R_+^3$. Its probability density function is

$$f_X(x) = \frac{1}{\alpha_1 \gamma_1} \left(1 + \frac{x - \alpha_2}{\alpha_1}\right)^{-(1+(1/\gamma_1))}, \quad x \geq \alpha_2. \tag{4.1}$$

Applying Theorem 3.1, one sees that the MLE of α_1, α_2 are determined by the real functions

$$S_1(x) = \frac{1 + \gamma_1}{1 + x}, \qquad L_2(x) = -\frac{1 + \gamma_1}{\gamma_1 x}. \tag{4.2}$$

According to Theorem 3.1, there are functions

$$q_1(x) = (1 + x)^{-(1+(\gamma_1/\gamma_1))}, \qquad q_2(x) = x^{-1+(\gamma_1/\gamma_1)}, \tag{4.3}$$

and constants $c_1 = -\gamma_1^{-1}, d_2 = -1$ such that

$$c_1 S_1(x) + 1 - c_1 = -x \cdot \frac{d}{dx} \ln\{q_1(x)\}, \qquad d_2 L_2(x) = -\frac{d}{dx} \ln\{q_2(x)\}, \tag{4.4}$$

and the compatibility condition (3.7) is fulfilled. For any random sample $\xi = (X_1, \ldots, X_n)$ from this family, one observes that the simultaneous maximum likelihood equations

$$\overline{\frac{1 + \gamma_1}{1 + ((\xi - \alpha_2)/\alpha_1)}} = 1, \qquad \overline{\frac{1}{1 + ((\xi - \alpha_2)/\alpha_1)}} = 0, \tag{4.5}$$

cannot have a common solution, hence the maximum likelihood method is not applicable.

The described pathological situation can be removed in a simple way thanks to Theorem 3.1. Our construction is motivated by the following question. What is the most general affine transform family with MLE of the affine parameter α_1 that is determined by the mean value equation $\overline{S_1((\xi-\alpha_2)/\alpha_1)} = 1$?. By Theorem 3.1, Case 1, there must exist a constant γ_2 and a function $q_1(x)$ such that

$$\gamma_2 S_1(x) + 1 - \gamma_2 = -x \cdot \frac{d}{dx} \ln\{q_1(x)\}. \tag{4.6}$$

Using [5], formula (3.1) one obtains

$$q_1(x) = x^{\gamma_2 - 1} \cdot \exp\left\{-\gamma_2 \int \frac{S_1(x)}{x} dx\right\} = x^{-(1+\gamma_1\gamma_2)} \cdot (1+x)^{(1+\gamma_1)\gamma_2}. \tag{4.7}$$

A corresponding probability density function is

$$f_X(x) = \frac{1}{C\alpha_1} \cdot \left(\frac{x-\alpha_2}{\alpha_1}\right)^{-(1+\gamma_1\gamma_2)} \cdot \left(1 + \frac{x-\alpha_2}{\alpha_1}\right)^{(1+\gamma_1)\gamma_2}, \quad x \geq \alpha_2. \tag{4.8}$$

One notes that two well-known subfamilies are included, namely, the *generalised Pareto* (4.1) obtained by setting $\gamma_1\gamma_2 = -1$, and the *Beta of type II* obtained by setting $p = -\gamma_1\gamma_2 > 0$, $q = -\gamma_2 > 0$. This suggests the name "generalised Pareto-Beta" but we prefer the simpler nomenclature "*Pareto type IV model*" for the new four-parameter family (4.8). Applying Theorem 3.1, one sees that the MLE of α_1 and α_2 are determined by

$$S_1(x) = \frac{1+\gamma_1}{1+x}, \qquad L_2(x) = \frac{(1+\gamma_1)\gamma_2}{x} - \frac{1+\gamma_1\gamma_2}{x-1}. \tag{4.9}$$

There are functions

$$q_1(x) = x^{-(1+\gamma_1\gamma_2)} \cdot (1+x)^{(1+\gamma_1)\gamma_2}, \qquad q_2(x) = (x-1)^{-(1+\gamma_1\gamma_2)} \cdot x^{(1+\gamma_1)\gamma_2}, \tag{4.10}$$

and constants $c_1 = \gamma_2, d_2 = -1$ such that

$$c_1 S_1(x) + 1 - c_1 = -x \cdot \frac{d}{dx} \ln\{q_1(x)\}, \quad d_2 L_2(x) = -\frac{d}{dx} \ln\{q_2(x)\}, \tag{4.11}$$

and the compatibility condition (3.7), that is,

$$\gamma_2 S_1(x-1) + 1 - \gamma_2 = -(x-1) L_2(x), \tag{4.12}$$

is fulfilled. For a random sample $\xi = (X_1, \dots, X_n)$, the MLE of α_1 and α_2 solves the simultaneous equations

$$\overline{\frac{1+\gamma_1}{1+((\xi-\alpha_2)/\alpha_1)}} = 1, \qquad \overline{\frac{1+\gamma_1\gamma_2}{(\xi-\alpha_2)/\alpha_1}} = \gamma_2. \tag{4.13}$$

The value of the normalising constant in (4.8) depends only on the shape vector $\gamma = (\gamma_1, \gamma_2)$.

Proposition 4.1. *Assume that $\gamma_2, \gamma_1\gamma_2$ are not integers. Then the normalising constant of the Pareto type IV model* (4.8) *is determined by the infinite series expansion*

$$C = C(\gamma_1, \gamma_2) = \sum_{k=0}^{\infty} \binom{(1+\gamma_1)\gamma_2}{k} \frac{2k - (1+\gamma_1)\gamma_2}{(k-\gamma_2)(k-\gamma_1\gamma_2)}, \tag{4.14}$$

where $\binom{\alpha}{k} = (\alpha(\alpha-1)\ldots(\alpha-k+1))/k!$, $k \geq 1$, $\binom{\alpha}{0} = 1$, *is a generalised binomial coefficient.*

Proof. From the observation made above, one notes that

$$C = \int_0^\infty q_1(x)dx = \int_0^\infty x^{-(1+\gamma_1\gamma_2)}(1+x)^{(1+\gamma_1)\gamma_2}dx = \int_0^\infty x^{\gamma_2-1}\left(1+x^{-1}\right)^{(1+\gamma_1)\gamma_2}dx. \tag{4.15}$$

To obtain convergent integrals, separate calculation in two parts and make a substitution to get

$$C = \int_0^1 x^{-(1+\gamma_1\gamma_2)}(1+x)^{(1+\gamma_1)\gamma_2}dx + \int_0^1 x^{-(1+\gamma_2)}(1+x)^{(1+\gamma_1)\gamma_2}dx. \tag{4.16}$$

The binomial expansion $(1+x)^\alpha = \sum_{k=0}^{\infty}\binom{\alpha}{k}x^k$, valid for $x \in (0,1)$ [14, (18.7), page 134], yields the series

$$C = \sum_{k=0}^{\infty}\binom{(1+\gamma_1)\gamma_2}{k} \cdot \left\{\int_0^1 x^{k-1-\gamma_1\gamma_2}dx + \int_0^1 x^{k-1-\gamma_2}dx\right\}. \tag{4.17}$$

Under the assumption $\gamma_2,\ \gamma_1\gamma_2 \neq k$, this implies without difficulty the expression (4.14). □

5. Conclusions and Outlook

The proposed method is not the only way to generalize the Pareto family (4.1). The recent note [9] extends this family to the family

$$f_X(x) = \frac{c}{\alpha_1\gamma_1} \cdot \left(\frac{x-\alpha_2}{\alpha_1}\right)^{c-1} \cdot \left(1 + \left(\frac{x-\alpha_2}{\alpha_1}\right)^c\right)^{-(1+(1/\gamma_1))}, \quad x \geq \alpha_2, \tag{5.1}$$

which looks similar to (4.8), except for the "power law" component in the second bracket, but has different statistical properties. An advantage of (5.1) is certainly the analytical closed-form expression for the survival function given by

$$S_X(x) = \left(1 + \left(\frac{x-\alpha_2}{\alpha_1}\right)^c\right)^{-(1+1/\gamma_1)}, \quad x \geq \alpha_2. \tag{5.2}$$

To conclude, several advantages of (4.8) can be noted, in particular, the simple MLE estimation of the affine parameters and the inclusion of the very important generalised Pareto distribution as a submodel. From a statistical viewpoint, the interest of the extended model (4.8) is two-fold. First, it may provide a better fit of the data than any submodel. Second, it yields a simple statistical procedure to choose among submodels like the generalised Pareto and the Beta of type II. Only the model “closest” to the full model will be retained. A detailed comparison of these two four parameter Pareto families is left to further research.

Acknowledgment

The author is grateful to the referees for careful reading of the manuscript and valuable comments.

References

[1] B. Efron, “Transformation theory: how normal is a family of distributions?” *The Annals of Statistics*, vol. 10, no. 2, pp. 323–339, 1982.

[2] W. Hürlimann, “General location transform of the order statistics from the exponential, Pareto and Weibull, with application to maximum likelihood estimation,” *Communications in Statistics: Theory and Methods*, vol. 29, no. 11, pp. 2535–2545, 2000.

[3] W. Hürlimann, “General affine transform families: why is the Pareto an exponential transform?” *Statistical Papers*, vol. 44, no. 4, pp. 499–519, 2003.

[4] E. J. Gumbel, *Statistics of Extremes*, Columbia University Press, New York, NY, USA, 1958.

[5] W. Hürlimann, “On the characterization of maximum likelihood estimators for location-scale families,” *Communications in Statistics: Theory and Methods*, vol. 27, no. 2, pp. 495–508, 1998.

[6] P. Embrechts, C. Klüppelberg, and Th. Mikosch, *Modelling Extremal Events for Insurance and Finance*, vol. 33 of *Applications of Mathematics*, Springer, Berlin, Germany, 1997.

[7] S. Kotz and S. Nadarajah, *Extreme Value Distributions: Theory and Applications*, Imperial College Press, London, UK, 2000.

[8] W. Hürlimann, “Higher-degree stop-loss transforms and stochastic orders (II) applications,” *Blätter der Deutschen Gesellschaft für Versicherungsmathematik*, vol. 24, no. 3, pp. 465–476, 2000.

[9] A. M. Abd Elfattah, E. A. Elsherpieny, and E. A. Hussein, “A new generalized Pareto distribution,” 2007, http://interstat.statjournals.net/YEAR/2007/abstracts/0712001.php.

[10] G. Bottazzi, “On the Pareto type III distribution,” Sant’Anna School of Advanced Studies, Pisa, Italy, 2007, http://www.lem.sssup.it/WPLem/files/2007-07.pdf.

[11] C. Kleiber and S. Kotz, *Statistical Size Distributions in Economics and Actuarial Sciences*, Wiley Series in Probability and Statistics, John Wiley & Sons, New York, NY, USA, 2003.

[12] A. K. Gupta and T. Varga, “An empirical estimation procedure,” *Metron*, vol. 52, no. 1-2, pp. 67–70, 1994.

[13] W. Walter, *Gewöhnliche Differentialgleichungen*, Eine Einführung. Heidelberger Taschenbücher, Band 110, Springer, Berlin, Germany, 1972.

[14] Ch. Blatter, *Analysis II*, Heidelberger Taschenbücher, Band 152, Springer, Berlin, Germany, 1974.