MARCEL DEKKER, INC. • 270 MADISON AVENUE • NEW YORK, NY 10016

COMMUNICATIONS IN STATISTICS
Theory and Methods
Vol. 32, No. 2, pp. 315–325, 2003

Information Matrix for Pareto(IV), Burr, and Related Distributions

Vytaras Brazauskas*

Department of Mathematical Sciences, University of Wisconsin-Milwaukee, Milwaukee, Wisconsin, USA

ABSTRACT

In this paper the exact form of information matrix for Pareto(IV) and related distributions is determined. The Pareto(IV) family being very general includes more specialized families of Pareto(I), Pareto(II), and Pareto(III), and the Burr family of distributions, as special cases. These distributions, for example, arise as tractable parametric models in actuarial science, economics, finance, and telecommunications. Additionally, a useful mathematical result with its own domain of importance is obtained. In particular, explicit formula for the improper integral $\int_0^\infty ((\log x)^m/(1+x)^{1+b})\,dx$, with $b > 0$ and non-negative integer m, is derived.

*Correspondence: Vytaras Brazauskas, Department of Mathematical Sciences, University of Wisconsin-Milwaukee, P.O. Box 413, Milwaukee, WI 53201-0413, USA; E-mail: vytaras@uwm.edu.

DOI: 10.1081/STA-120018188
0361-0926 (Print); 1532-415X (Online)

Key Words: Gamma function; Information; Pareto models; Polygamma functions.

1. INTRODUCTION

In this paper the exact form of (Fisher) information matrix for Pareto(IV) and related distributions is determined. It is well-known that information matrix serves as a valuable tool for derivation of covariance matrix in the asymptotic distribution of maximum likelihood estimators (MLE). As discussed in Serfling (1980), Sec. 4.1, under suitable regularity conditions, the determinant (divided by the sample size) of the asymptotic covariance matrix of MLE reaches an optimal lower bound for the volume of the "spread ellipsoid" of joint estimators. In the univariate case, this optimality property of MLE is widely used in the "robustness versus efficiency" studies as a quantitative benchmark for efficiency considerations. See, for example, Brazauskas and Serfling (2000a, 2000b), Hampel et al. (1986), Huber (1981), Kimber (1983a, 1983b), and Lehmann (1983), Chapter 5.

The Pareto(IV) family being very general includes more specialized families of Pareto(I), Pareto(II), and Pareto(III), and the Burr family of distributions, as special cases. These distributions are suitable for situations involving relatively high probability in the upper tail. More specifically, such models have been formulated in the context of actuarial science, economics, finance, and teletraffic, for example, for distributions of variables such as sizes of insurance claims, sizes of firms, incomes in a population of people, stock price fluctuations, and length of telephone calls. (See Arnold (1983) and Johnson et al. (1994), Chapter 19, for a broad discussion of Pareto models, and their diverse applications.) New application contexts continue to be found. For example, Crato et al. (1997) have recently discovered Pareto-type tail behavior in the cost distributions of combinatorial search algorithms.

As a by-product of the main result, a useful mathematical formula (which plays a very important role in our derivations) is obtained. In particular, explicit formula for the improper integral

$$\int_0^\infty \frac{(\log x)^m}{(1+x)^{1+b}}\, dx, \tag{1.1}$$

with $b > 0$ and non-negative integer m, is derived. Surprisingly, however, computation of integral (1.1) has not been addressed in the mathematical literature including such comprehensive handbooks of mathematics as Abramowitz and Stegun (1972) and Harris and Stocker (1998).

A hierarchy of Pareto models as well as their relation to the Burr family is discussed in Sec. 2. Elements of the information matrix for Pareto(IV), Burr, and related distributions are computed in Sec. 3. Derivation of integral (1.1) along with other intermediate integrals and formulas are presented in the Appendix.

2. PARETO(IV) AND RELATED DISTRIBUTIONS

As discussed in Arnold (1983), Chapter 3, a hierarchy of Pareto distributions is established by starting with the classical Pareto(I) distribution and subsequently introducing additional parameters which relate to *location*, *scale*, *shape*, and *inequality*. Such an approach leads to a very general family of distributions, called the Pareto(IV) family, with the cumulative distribution function

$$F(x) = 1 - \left[1 + \left(\frac{x-\mu}{\sigma}\right)^{1/\gamma}\right]^{-\alpha}, \quad x > \mu, \tag{2.1}$$

where $-\infty < \mu < +\infty$ is the location parameter, $\sigma > 0$ is the scale parameter, $\gamma > 0$ is the inequality parameter, and $\alpha > 0$ is the shape parameter which characterizes the tail of the distribution. We denote this distribution by Pareto(IV) $(\mu, \sigma, \gamma, \alpha)$.

Note that in general statistical science there is no such type of parameters like "inequality." Parameter γ is called the inequality parameter because of its interpretation in the economics context. That is, if we choose $\alpha = 1$ and $\mu = 0$ in expression (2.1), then parameter γ $(\gamma \leq 1)$ is precisely the Gini index of inequality.

Clearly, the other three types of Pareto distributions can be identified as special cases of the Pareto(IV) family by appropriately choosing parameters in Eq. (2.1) (see Arnold (1983), pp. 44–45):

$$\begin{aligned} \text{Pareto(I)}\,(\sigma,\, \alpha) &= \text{Pareto(IV)}\ (\sigma,\, \sigma,\, 1,\, \alpha), \\ \text{Pareto(II)}\ (\mu,\, \sigma,\, \alpha) &= \text{Pareto(IV)}\ (\mu,\, \sigma,\, 1,\, \alpha), \\ \text{Pareto(III)}\ (\mu,\, \sigma,\, \gamma) &= \text{Pareto(IV)}\ (\mu,\, \sigma,\, \gamma,\, 1). \end{aligned}$$

The Burr family of distributions is also sufficiently flexible and enjoys long popularity in the actuarial science literature (see, for example, Daykin et al. (1994) and Klugman et al. (1998)). However, even this family can be treated as a special case of Pareto(IV):

$$\text{Burr}\,(\sigma,\,\gamma,\,\alpha) = \text{Pareto(IV)}\,(0,\,\sigma,\,1/\gamma,\,\alpha)$$

(see Klugman et al. (1998), p. 574).

In order to make the Pareto(IV) distribution a regular family, we assume that parameter μ is known and, without loss of generality, equal to 0. (For regularity conditions on F and their interpretation, see, for example, Serfling (1980), pp. 144–145.) Also, we note that this assumption is not too restrictive for modeling purposes because in typical applications the lower limit of variables of interest is known. In insurance and reinsurance context, for example, the lower limit is pre-defined by a contract and can be represented as a deductible or a retention level. (This perhaps is one of the main reasons why generality of the Burr distribution remains sufficient for actuarial modeling.)

3. INFORMATION MATRIX FOR PARETO(IV)

Suppose X is a random variable with the probability density function $f_\Theta(\cdot)$ where $\Theta = (\theta_1, \ldots, \theta_k)$. Then the information matrix $\mathbf{I}(\Theta)$ is the $k \times k$ matrix with elements

$$I_{ij}(\Theta) = \mathbf{E}_\Theta\left[\frac{\partial \log f_\Theta(X)}{\partial\theta_i} \cdot \frac{\partial \log f_\Theta(X)}{\partial\theta_j}\right].$$

For the Pareto(IV) $(0,\,\sigma,\,\gamma,\,\alpha)$ distribution, we have $\Theta = (\theta_1,\,\theta_2,\,\theta_3) = (\sigma,\,\gamma,\alpha)$ and the density function

$$f(x) = \frac{\alpha}{\gamma\sigma} \cdot \frac{(x/\sigma)^{1/\gamma-1}}{(1+(x/\sigma)^{1/\gamma})^{\alpha+1}}. \tag{3.1}$$

Thus the required partial derivatives are

$$\begin{aligned}
\frac{\partial \log f_\Theta(x)}{\partial\theta_1} &= \frac{\partial \log f(x)}{\partial\sigma} = \frac{\alpha}{\gamma\sigma} - \frac{\alpha+1}{\gamma\sigma} \cdot \frac{1}{1+(x/\sigma)^{1/\gamma}},\\
\frac{\partial \log f_\Theta(x)}{\partial\theta_2} &= \frac{\partial \log f(x)}{\partial\gamma} = \frac{\alpha}{\gamma^2}\log\left(\frac{x}{\sigma}\right) - \frac{\alpha+1}{\gamma^2} \cdot \frac{\log(x/\sigma)}{1+(x/\sigma)^{1/\gamma}} - \frac{1}{\gamma},\\
\frac{\partial \log f_\Theta(x)}{\partial\theta_3} &= \frac{\partial \log f(x)}{\partial\alpha} = -\log\left(1+\left(\frac{x}{\sigma}\right)^{1/\gamma}\right) + \frac{1}{\alpha}.
\end{aligned}$$

Note that for this problem the first two partial derivatives $\partial^2/\partial\theta_i\partial\theta_j$ of $f_\Theta(\cdot)$ exist and, therefore, an alternative (based on the second-order derivatives) formula for computation of information matrix elements may be used. We found, however, that, except for a couple of cases, this approach does not simplify our derivations. Therefore it is not pursued here.

Since the information matrix $\mathbf{I}(\Theta)$ is symmetric, it is enough to find elements $I_{11}(\Theta)$, $I_{12}(\Theta)$, $I_{13}(\Theta)$, $I_{22}(\Theta)$, $I_{23}(\Theta)$, and $I_{33}(\Theta)$. Derivation of these elements is based on the following strategy: first, we express each $I_{ij}(\Theta)$, $1 \le i \le j \le 3$, in terms of integrals $A1$–$A5$, $B1$–$B4$, $C1$, $C2$, $D1$, $D2$, which are defined (and their explicit formulas are presented) in Appendix A.2; then, tedious algebraic simplifications yield the following formulas.

$$\begin{aligned}
I_{11}(\Theta) &= \int_0^\infty \left[\frac{\partial \log f(x)}{\partial\sigma}\right]^2 f(x)\,dx \\
&= \left(\frac{\alpha}{\gamma\sigma}\right)^2 - \frac{2\alpha(\alpha+1)}{(\gamma\sigma)^2}\cdot A1 + \left(\frac{\alpha+1}{\gamma\sigma}\right)^2\cdot A2 = \frac{\alpha}{(\gamma\sigma)^2(\alpha+2)}, \\
I_{12}(\Theta) &= \int_0^\infty \left[\frac{\partial \log f(x)}{\partial\sigma}\cdot\frac{\partial \log f(x)}{\partial\gamma}\right] f(x)\,dx \\
&= \frac{\alpha^2}{\gamma^3\sigma}\cdot B1 - \frac{2\alpha(\alpha+1)}{\gamma^3\sigma}\cdot B2 - \frac{\alpha}{\gamma^2\sigma}\frac{(\alpha+1)^2}{\gamma^3\sigma}\cdot B3 + \frac{\alpha+1}{\gamma^2\sigma}\cdot A1 \\
&= \frac{\alpha}{\gamma^2\sigma(\alpha+2)}\left(\Gamma'(1) - \frac{\Gamma'(\alpha)}{\Gamma(\alpha)} + \frac{\alpha-1}{\alpha}\right), \\
I_{13}(\Theta) &= \int_0^\infty \left[\frac{\partial \log f(x)}{\partial\sigma}\cdot\frac{\partial \log f(x)}{\partial\alpha}\right] f(x)\,dx \\
&= \frac{1}{\gamma\sigma} - \frac{\alpha}{\gamma\sigma}\cdot A3 - \frac{\alpha+1}{\gamma\sigma\alpha}\cdot A1 + \frac{\alpha+1}{\gamma\sigma}\cdot A5 = -\frac{1}{\gamma\sigma(\alpha+1)}, \\
I_{22}(\Theta) &= \int_0^\infty \left[\frac{\partial \log f(x)}{\partial\gamma}\right]^2 f(x)\,dx = \frac{\alpha^2}{\gamma^4}\cdot B4 + \frac{(\alpha+1)^2}{\gamma^4}\cdot C2 + \frac{1}{\gamma^2} \\
&\quad - \frac{2\alpha(\alpha+1)}{\gamma^4}\cdot C1 - \frac{2\alpha}{\gamma^3}\cdot B1 + \frac{2(\alpha+1)}{\gamma^3}\cdot B2 \\
&= \frac{\alpha}{\gamma^2(\alpha+2)}\left(\frac{\Gamma''(\alpha)}{\Gamma(\alpha)} + \Gamma''(1) + 1\right) + \frac{2(\alpha-1)}{\gamma^2(\alpha+2)}\left(\Gamma'(1) - \frac{\Gamma'(\alpha)}{\Gamma(\alpha)}\right) \\
&\quad - \frac{2\alpha}{\gamma^2(\alpha+2)}\cdot\frac{\Gamma'(\alpha)}{\Gamma(\alpha)}\cdot\Gamma'(1),
\end{aligned}$$

$$
\begin{aligned}
I_{23}(\Theta) &= \int_0^{\infty} \left[\frac{\partial \log f(x)}{\partial \gamma} \cdot \frac{\partial \log f(x)}{\partial \alpha} \right] f(x)\, dx \\
&= \frac{1}{\gamma^2} \cdot B1 - \frac{\alpha+1}{\alpha\gamma^2} \cdot B2 - \frac{1}{\alpha\gamma} - \frac{\alpha}{\gamma^2} \cdot D1 + \frac{\alpha+1}{\gamma^2} \cdot D2 + \frac{1}{\gamma} \cdot A3 \\
&= \frac{1}{\gamma(\alpha+1)} \left(\frac{\Gamma'(\alpha)}{\Gamma(\alpha)} - \Gamma'(1) - 1 \right), \\
I_{33}(\Theta) &= \int_0^{\infty} \left[\frac{\partial \log f(x)}{\partial \alpha} \right]^2 f(x)\, dx = \frac{1}{\alpha^2} - \frac{2}{\alpha} \cdot A3 + A4 = \frac{1}{\alpha^2}.
\end{aligned}
$$

Finally, elements $I_{12}(\Theta)$, $I_{22}(\Theta)$, and $I_{23}(\Theta)$ can be written in a simpler form by using the polygamma functions $\psi^{(n)}(a) = (d^n/da^n)(\Gamma'(a)/\Gamma(a))$, for $a > 0$ and integer $n \geq 0$. Specifically, we use digamma $\psi(a) = \psi^{(0)}(a)$ and trigamma $\psi'(a)$ functions. Thus the information matrix, $\mathbf{I}_P(\Theta)$, for the Pareto(IV) (0, σ, γ, α) distribution is given by:

$$
\begin{pmatrix}
\dfrac{\alpha}{(\gamma\sigma)^2(\alpha+2)} & \dfrac{\alpha[\psi(1)-\psi(\alpha)+1]-1}{\gamma^2\sigma(\alpha+2)} & -\dfrac{1}{\gamma\sigma(\alpha+1)} \\
\dfrac{\alpha[\psi(1)-\psi(\alpha)+1]-1}{\gamma^2\sigma(\alpha+2)} & \dfrac{\alpha\left[(\psi(\alpha)-\psi(1)-1)^2+\psi'(\alpha)+\psi'(1)\right]+2(\psi(\alpha)-\psi(1))}{\gamma^2(\alpha+2)} & \dfrac{\psi(\alpha)-\psi(1)-1}{\gamma(\alpha+1)} \\
-\dfrac{1}{\gamma\sigma(\alpha+1)} & \dfrac{\psi(\alpha)-\psi(1)-1}{\gamma(\alpha+1)} & \dfrac{1}{\alpha^2}
\end{pmatrix}
$$

3.1. Special Cases

3.1.1. Burr(σ, γ, α) Distribution

Since the Burr distribution is a reparametrization of Pareto(IV) (0, σ, γ, α), it follows from Lehmann (1983), Sec. 2.7, that its information matrix $\mathbf{I}_B(\Theta)$ can be derived from $J\,\mathbf{I}_P(\Theta)\,J'$, where J is the Jacobian matrix of the transformation of variables. Thus $\mathbf{I}_B(\Theta)$ is then given by:

$$
\begin{pmatrix}
\dfrac{\alpha}{(\gamma\sigma)^2(\alpha+2)} & \dfrac{\alpha[\psi(\alpha)-\psi(1)-1]+1}{\sigma(\alpha+2)} & -\dfrac{1}{\gamma\sigma(\alpha+1)} \\
\dfrac{\alpha[\psi(\alpha)-\psi(1)-1]+1}{\sigma(\alpha+2)} & \dfrac{\gamma^2\left[\alpha\left[(\psi(\alpha)-\psi(1)-1)^2+\psi'(\alpha)+\psi'(1)\right]+2(\psi(\alpha)-\psi(1))\right]}{\alpha+2} & \dfrac{\gamma[\psi(1)-\psi(\alpha)+1]}{\alpha+1} \\
-\dfrac{1}{\gamma\sigma(\alpha+1)} & \dfrac{\gamma[\psi(1)-\psi(\alpha)+1]}{\alpha+1} & \dfrac{1}{\alpha^2}
\end{pmatrix}
$$

3.1.2. Pareto(III) $(0, \sigma, \gamma)$ Distribution

This is a special case of Pareto(IV) with $\alpha = 1$. Therefore third row and third column (these represent information about parameter α) in $\mathbf{I}_P(\Theta)$ vanish. And into expressions of the remaining elements we substitute $\alpha = 1$. This yields

$$\mathbf{I}^*(\sigma, \gamma) = \begin{pmatrix} \frac{1}{3(\gamma\sigma)^2} & 0 \\ 0 & \frac{1 + 2\psi'(1)}{3\gamma^2} \end{pmatrix} = \begin{pmatrix} \frac{1}{3(\gamma\sigma)^2} & 0 \\ 0 & \frac{\pi^2 + 3}{9\gamma^2} \end{pmatrix}$$

3.1.3. Pareto(II) $(0, \sigma, \alpha)$ Distribution

This is a special case of Pareto(IV) with $\gamma = 1$. Therefore second row and second column (these represent information about parameter γ) in $\mathbf{I}_P(\Theta)$ vanish. And into the remaining formulas we substitute $\gamma = 1$. This yields

$$\mathbf{I}^{**}(\sigma, \alpha) = \begin{pmatrix} \frac{\alpha}{\sigma^2(\alpha + 2)} & -\frac{1}{\sigma(\alpha + 1)} \\ -\frac{1}{\sigma(\alpha + 1)} & \frac{1}{\alpha^2} \end{pmatrix}$$

It should be noted here that information matrices $\mathbf{I}^*(\sigma, \gamma)$ and $\mathbf{I}^{**}(\sigma, \alpha)$ are readily available in Arnold (1983), p. 210.

APPENDIX

A.1. Three Lemmas

Lemma 1. *For $b > 0$ and non-negative integer m,*

$$\int_0^\infty \frac{(\log x)^m}{(1 + x)^{1+b}} \, dx = \frac{1}{\Gamma(1 + b)} \sum_{i=0}^{m} \binom{m}{i} (-1)^{m-i} \, \Gamma^{(i)}(1) \, \Gamma^{(m-i)}(b), \tag{A.1}$$

where $\Gamma(a)=\int_0^\infty x^{a-1}e^{-x}\,dx$ *and* $\Gamma^{(n)}(a)=\int_0^\infty(\log x)^n\,x^{a-1}\,e^{-x}\,dx$ *are gamma function and its n-th order derivative, respectively.*

Proof. We start with the following fact (Abramowitz and Stegun (1972), p. 255):

$$\frac{1}{x^a}=\frac{1}{\Gamma(a)}\int_0^\infty t^{a-1}e^{-x\,t}\,dt, \quad x>0, \quad a>0.$$

Next, after applying the above formula to the integral in Eq. (A.1) and interchanging the order of integration we have

$$\int_0^\infty \frac{(\log x)^m}{(1+x)^{1+b}}\,dx=\frac{1}{\Gamma(1+b)}\int_0^\infty t^b\,e^{-t}\left[\int_0^\infty e^{-t\,x}\,(\log x)^m\,dx\right]dt.$$

Further, substitution of variables $z=tx$ and application of the binomial formula to $(\log z-\log t)^m$ lead to the following expression

$$\frac{1}{\Gamma(1+b)}\int_0^\infty t^{\,b-1}\,e^{-t}\int_0^\infty\left[\sum_{i=0}^{m}\binom{m}{i}(\log z)^i\,(-\log t)^{m-i}\right]e^{-z}\,dz\,dt.$$

Finally, simple reorganization of terms yields the result. ◁

Lemma 2. *For* $b>1$ *and integer* $n\geq 1$,

$$\Gamma^{(n)}(b)=(b-1)\Gamma^{(n)}(b-1)+n\Gamma^{(n-1)}(b-1), \tag{A.2}$$

where $\Gamma(\cdot)$ *and* $\Gamma^{(n)}(\cdot)$ *denote gamma function and its n-th order derivative, respectively.*

Proof. Integration by parts of the first derivative of the gamma function, $\Gamma'(b)$, leads to

$$\Gamma'(b)=(b-1)\Gamma'(b-1)+\Gamma(b-1).$$

Differentiation of this equation $n-1$ times yields Eq. (A.2). ◁

Lemma 3. *For* $b>0$,

$$\int_0^\infty \frac{\log(1+x)\,\log x}{(1+x)^{1+b}}\,dx=\frac{1}{b}\left(\frac{\Gamma''(b)}{\Gamma(b)}-\left[\frac{\Gamma'(b)}{\Gamma(b)}\right]^2-\frac{1}{b}\left[\frac{\Gamma'(b)}{\Gamma(b)}-\Gamma'(1)\right]\right), \tag{A.3}$$

where $\Gamma(\cdot)$ *and* $\Gamma^{(n)}(\cdot)$ *denote gamma function and its n-th order derivative, respectively.*

Proof. Note that

$$\int_0^\infty \frac{\log(1+x)\log x}{(1+x)^{1+b}}\,dx = \frac{d}{d\,b}\left[-\int_0^\infty \frac{\log x}{(1+x)^{1+b}}\,dx\right].$$

Formula (A.1) and differentiation of the right-hand side yield (A.3). ◁

A.2. Useful Integrals

In all expressions below function $f(x)$ is the Pareto(IV) $(0, \sigma, \gamma, \alpha)$ density function given by Eq. (3.1).

Integrals $A1-A5$ are derived via straightforward integration.

$$A1 = \int_0^\infty \frac{f(x)}{1+(x/\sigma)^{1/\gamma}}\,dx = \frac{\alpha}{\alpha+1},$$

$$A2 = \int_0^\infty \frac{f(x)}{\left[1+(x/\sigma)^{1/\gamma}\right]^2}\,dx = \frac{\alpha}{\alpha+2},$$

$$A3 = \int_0^\infty \log\left[1+(x/\sigma)^{1/\gamma}\right] f(x)\,dx = \frac{1}{\alpha},$$

$$A4 = \int_0^\infty \log^2\left[1+(x/\sigma)^{1/\gamma}\right] f(x)\,dx = \frac{2}{\alpha^2},$$

$$A5 = \int_0^\infty \frac{\log\left[1+(x/\sigma)^{1/\gamma}\right]}{1+(x/\sigma)^{1/\gamma}} f(x)\,dx = \frac{\alpha}{(\alpha+1)^2}.$$

Integrals $B1$–$B4$ are computed by making use of formula (A.1).

$$B1 = \int_0^\infty \log(x/\sigma) f(x)\,dx = \gamma\left(\Gamma'(1) - \frac{\Gamma'(\alpha)}{\Gamma(\alpha)}\right),$$

$$B2 = \int_0^\infty \frac{\log(x/\sigma)}{1+(x/\sigma)^{1/\gamma}} f(x)\,dx = \frac{\alpha\gamma}{\alpha+1}\left(\Gamma'(1) - \frac{\Gamma'(\alpha)}{\Gamma(\alpha)} - \frac{1}{\alpha}\right),$$

$$B3 = \int_0^\infty \frac{\log(x/\sigma)}{(1+(x/\sigma)^{1/\gamma})^2} f(x)\,dx = \frac{\alpha\gamma}{\alpha+2}\left(\Gamma'(1) - \frac{\Gamma'(\alpha)}{\Gamma(\alpha)} - \frac{1}{\alpha} - \frac{1}{\alpha+1}\right),$$

$$B4 = \int_0^\infty \log^2(x/\sigma) f(x)\,dx = \gamma^2\left(\Gamma''(1) - 2\Gamma'(1)\frac{\Gamma'(\alpha)}{\Gamma(\alpha)} + \frac{\Gamma''(\alpha)}{\Gamma(\alpha)}\right).$$

Lemmas 1 and 2 are applied in computation of integrals $C1$ and $C2$.

$$C1 = \int_0^\infty \frac{\log^2(x/\sigma)}{1+(x/\sigma)^{1/\gamma}} f(x)\,dx$$
$$= \frac{\alpha\gamma^2}{\alpha+1}\left(\Gamma''(1) - \frac{2\Gamma'(\alpha)}{\Gamma(\alpha)}\left[\frac{1}{\alpha} - \Gamma'(1)\right] + \frac{\Gamma''(\alpha)}{\Gamma(\alpha)} - \frac{2\Gamma'(1)}{\alpha}\right),$$
$$C2 = \int_0^\infty \frac{\log^2(x/\sigma)}{(1+(x/\sigma)^{1/\gamma})^2} f(x)\,dx = \frac{\alpha\gamma^2}{\alpha+2}\left(\Gamma''(1) - 2\Gamma'(1)\left[\frac{1}{\alpha} + \frac{1}{\alpha+1}\right]\right.$$
$$\left. + \frac{\Gamma''(\alpha)}{\Gamma(\alpha)} + \frac{2}{\alpha(\alpha+1)} - \frac{2\Gamma'(\alpha)}{\Gamma(\alpha)}\left[\frac{1}{\alpha} + \frac{1}{\alpha+1} - \Gamma'(1)\right]\right).$$

Integrals $D1$ and $D2$ are derived by employing Lemmas 1–3.

$$D1 = \int_0^\infty \log(x/\sigma)\log(1+(x/\sigma)^{1/\gamma}) f(x)\,dx$$
$$= \gamma\left(\frac{\Gamma''(\alpha)}{\Gamma(\alpha)} - \left[\frac{\Gamma'(\alpha)}{\Gamma(\alpha)}\right]^2 - \frac{1}{\alpha}\left[\frac{\Gamma'(\alpha)}{\Gamma(\alpha)} - \Gamma'(1)\right]\right),$$
$$D2 = \int_0^\infty \frac{\log(x/\sigma)\log(1+(x/\sigma)^{1/\gamma})}{1+(x/\sigma)^{1/\gamma}} f(x)\,dx$$
$$= \frac{\alpha\gamma}{\alpha+1}\left(\frac{\Gamma''(\alpha)}{\Gamma(\alpha)} - \left[\frac{\Gamma'(\alpha)}{\Gamma(\alpha)}\right]^2 - \frac{1}{\alpha+1}\left[\frac{\Gamma'(\alpha)}{\Gamma(\alpha)} - \Gamma'(1)\right] - \frac{2\alpha+1}{\alpha^2(\alpha+1)}\right).$$

REFERENCES

Abramowitz, M., Stegun, I. A. (1972). *Handbook of Mathematical Functions*. National Bureau of Standards, Applied Mathematics Series, No. 55.

Arnold, B. C. (1983). *Pareto Distributions*. Fairland, Maryland: International Cooperative Publishing House.

Brazauskas, V., Serfling, R. (2000a). Robust estimation of tail parameters for two-parameter Pareto and exponential models via generalized quantile statistics. *Extremes* 3(3):231–249.

Brazauskas, V., Serfling, R. (2000b). Robust and efficient estimation of the tail index of a single-parameter Pareto distribution. *North American Actuarial Journal* 4(4):12–27.

Daykin, C. D., Pentikäinen, T., Pesonen, M. (1994). *Practical Risk Theory for Actuaries*. London: Chapman and Hall.

Gomes, C. P., Selman, B., Crato, N. (1997). Heavy-tailed distributions in combinatorial search. In: Smolka, G., ed. *Principles and Practice of Constraint Programming CP-97*. Lecture Notes in Computer Science. Vol. 1330, 121–135.

Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J., Stahel, W. A. (1986). *Robust Statistics: The Approach Based On Influence Functions*. New York: Wiley.

Harris, J. W., Stocker, H. (1998). *Handbook of Mathematics and Computational Science*. New York: Springer.

Huber, P. J. (1981). *Robust Statistics*. New York: Wiley.

Johnson, N. L., Kotz, S., Balakrishnan, N. (1994). *Continuous Univariate Distributions*. Vol. 1. 2nd ed. New York: Wiley.

Kimber, A. C. (1983a). Trimming in gamma samples. *Applied Statistics* 32(1):7–14.

Kimber, A. C. (1983b). Comparison of some robust estimators of scale in gamma samples with known shape. *Journal of Statistical Computation and Simulation* 18:273–286.

Klugman, S. A., Panjer, H. H., Willmot, G. E. (1998). *Loss Models: From Data to Decisions*. New York: Wiley.

Lehmann, E. L. (1983). *Theory of Point Estimation*. New York: Wiley.

Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics*. New York: Wiley.