Laboratory of Economics and Management
Sant'Anna School of Advanced Studies
Piazza Martiri della Libertà, 33 - 56127 PISA (Italy)
Tel. +39-050-883-343 Fax +39-050-883-344
Email: lem@sssup.it Web Page: http://www.lem.sssup.it/

LEM

Working Paper Series

On the Pareto Type III distribution

Giulio Bottazzi[§]

[§] LEM-Scuola Superiore Sant'Anna, Pisa, Italy

2007/07 June 2009

ISSN (online) 2284-0400

On the Pareto Type III distribution*

Giulio Bottazzi†

Scuola Superiore S.Anna

July 2, 2009

Abstract

This note analyzes the distributional properties of Pareto Type III random variables. The orignal two parameters distribution proposed by Pareto is expanded in a three parameters version and both its density and characteristic function are derived. The analytic expression of the inverse distribution function is also obtained, together with a simple series expansion of its moments of any order. Finally, we propose a simple statistical exercise designed to show the increased reliability of the Pareto Type III distribution in describing asymptotically dumped power-like behaviors.

Keywords: Pareto Distribution, Fat Tails, Power Law Distribution, Zipf Law

*I would like to thank Herman Rubin for his suggestion of using Lambert function to express the Pareto Type III inverse distribution function. I also thank Federico Tamagni and Angelo Secchi for a number of helpful comments. Of course, all the mistakes remain mine. Support from the Scuola Superiore Sant'Anna (grant E6006GB) and from the EU (Contract No 12410 (NEST)) is gratefully acknowledged.

†Scuola Superiore Sant'Anna, P.za Martiri della Liberta' 33, 56127 Pisa, Italy. *E-mail*: bottazzi@sssup.it. *Phone*: +39-050-883343. *Fax*: +39-050-883344.

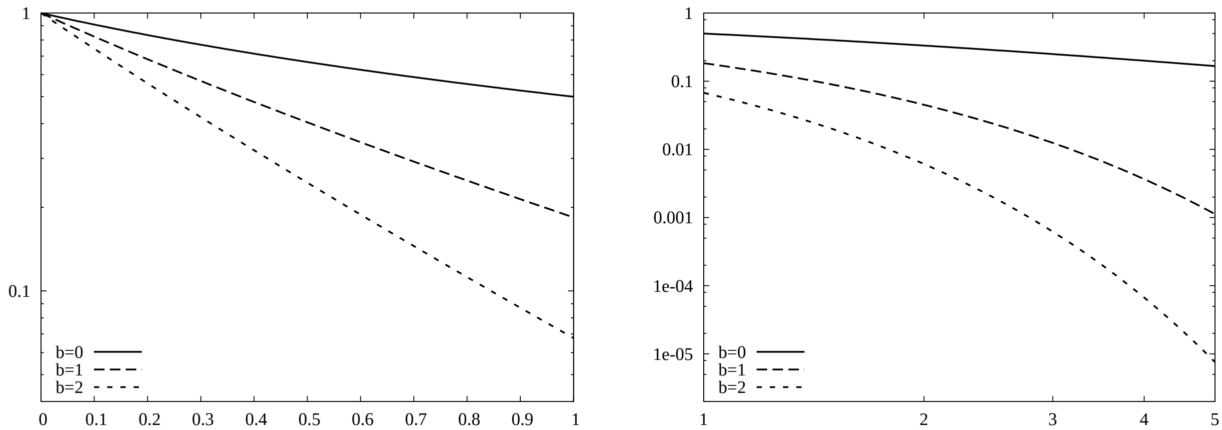

Figure 1: Lower (left) and upper (right) tail of the survival function of the Pareto III distribution for $a = 1$, $s = 1$ and different values of the exponential coefficient b. The case $b = 0$ corresponds to the Zipf's law $\log(1 - F(x)) \sim x^{-1}$.

1 Introduction

Zipf's Law (Zipf, 1949), and Pareto Laws in general (Pareto, 1897), have been assumed, since a relative long time, as reliable descriptions of the distributional properties of variables characterizing different social and natural phenomena (see Newman (2005), the critical review in Kleiber and Kotz (2003) and references therein). These Laws are usually assumed valid when sufficiently large realizations of the variable of interest are considered, that is for the upper tail of the associated distributions. If a random variable $\mathbf{x}$ follows the Pareto Law with tail index a, the associated survival function possesses a power-like asymptotic behaviour, $1-F(x) \sim x^{-a}$, so that its n-th moment $\mathrm{E}[\mathbf{x}^n]$, for sufficiently large values of n, diverges. This property is particularly relevant when small values of the tail parameter are considered. For instance, when $a \leq 2$, the variance (n = 2) is already absent. Contrary to the Pareto hypothesis, however, the empirical evidence often suggests finite values for the central moments, also when heavy-tail distributed variables are considered. Sometimes this can be understood with the presence of a natural upper bound which limits the largest attainable value of a given variable. Hence, in several cases it has been reported that a truncated Pareto provides a better fit to the data (Burroughs and Tebbens , 2001) than a pure power Law. In this paper I introduce a new three parameters version of the two parameters family of distributions originally proposed in Pareto (1896). The added parameter is a "dispersion" parameter which explicitly accounts for the typical scale of the underlying random variable, thus reducing the negative effect of the latter on the estimation of the power and exponential coefficients reported in Creedy (1977). Since the shape of the distribution is essentially the same of the orignal two parameters family proposed by Pareto, following Kleiber and Kotz (2003) I retain the name of Pareto Type III distribution. The exponential dumping in the upper tail makes this distribution particularly suitable in describing all those samples which display power behavior for intermediate values and a more than power-like decrease in probability above a certain threshold or for particularly large observations (c.f. the examples in Burroughs and Tebbens (2001)).

2 The Pareto Type III Distribution

Consider the three parameters family of distributions

$$F(x) = 1 - \left(\frac{x}{s} + 1\right)^{-a} e^{-b\frac{x}{s}} \tag{1}$$

where $s > 0$ is a scale parameter, $a > 0$ the "power" coefficient and ≥ 0 the "exponential" coefficient.

Notice that when the exponential coefficient is set to zero, $b = 0$, the previous equation reduces to a Pareto Type II distribution (Kleiber and Kotz, 2003), which is asymptotically equivalent to the original Pareto distribution. As discussed above, with respect to the original formulation of Pareto and some later work, the functional expression proposed in (1) contains one more parameter, the "scale" parameter s. The associated density reads

$$f(x) = \frac{1}{s} \left(\frac{x}{s} + 1\right)^{-a} e^{-b\frac{x}{s}} \left(b + \frac{a}{x/s + 1}\right) . \tag{2}$$

It is immediate to see, with direct derivation, that $f'(x) > 0$ for any allowed value of the parameters, so that the density in (2) is unimodal with mode in $x = 0$.

The effect of the values of the exponential coefficient b on the overall shape of the distribution can be see in Fig. 1, where the survival function $1 - F(x)$ is reported. For $x/s << 1$ (left panel), when b is large the exponential factor becomes leading and the survival function decreases linearly on the normal x scale. Conversely, when $x/s >> 1$ (right panel) an increase in b introduce a deviation (on the log-log graph) from the stright line characteristic of the Pareto power like behavior.

2.1 Characteristic Function and Moments

Following Abramowitz and Stegun (1964) (equation 6.5.20, p. 262)) define

$$\alpha_n(x) = \int_1^{+\infty} dt\, e^{-xt}\, t^n = x^{-n-1}\, \Gamma(n+1, x) \tag{3}$$

where $\Gamma(n, x)$ stands for the incomplete gamma function

$$\Gamma(n, x) = \int_x^{+\infty} dt\, e^{-t}\, t^{n-1} .$$

Using the properties of $\Gamma(n, x)$, it is immediate to obtain the following recurrence relation

$$\alpha_n(x) = \frac{e^{-x}}{x} + \frac{n}{x}\alpha_{n-1}(x) . \tag{4}$$

Using this relation it is easy to show that

Theorem 2.1 *The characteristic function $\phi(k) = E[e^{ik\mathbf{x}}]$ where x is distributed according to* (1) *reads*

$$\phi(k) = 1 + e^{-iks+b}\, iks\, \alpha_{-a}(b - iks) . \tag{5}$$

Proof. From the definition above and using (2) one has

$$\phi(k) = \int_0^{+\infty} dx\, e^{ikx} \frac{1}{s} \left(\frac{x}{s} + 1\right)^{-a} e^{-b\frac{x}{s}} \left(b + \frac{a}{x/s + 1}\right)$$

which, after the change of variable $z = x/s + 1$, reduces to

$$\phi(k) = \int_1^{+\infty} dz\, z^{-a} e^{-(b-iks)z} \left(b + \frac{a}{z}\right) .$$

Using (3) this can be rewritten as

$$\phi(k) = e^{-iks+b} \left(b\alpha_{-a}(b - iks) + a\alpha_{-a-1}(b - iks)\right)$$

which using (4) reduces to (5).

Q.E.D.

The moments of the Pareto III distribution can in principle be obtained from the characteristic function defined in (5). For instance, one immediately has that the mean value m_1 is

$$m_1 = (-i)\frac{d}{dk}\bigg|_{k=0} \phi(k) = 1 + e^b \alpha_{-a}(b)\ . \tag{6}$$

For higher moments, however, can be more practical to directly derive a series expansion. The following applies

Theorem 2.2 *The moments $m_n = E[\mathbf{x}^n]$ where x is distributed according to* (1) *admit the following representation:*

$$m_n = s^n\, e^b\, b^a \sum_{h=0}^{n} \binom{n}{h} (-1)^{n-h} \frac{h}{b^h}\, \Gamma(h-a,b)\ . \tag{7}$$

Proof. From the definition

$$m_n = \int_0^{+\infty} dx x^n \frac{1}{s}\left(\frac{x}{s}+1\right)^{-a} e^{-b\frac{x}{s}} \left(b + \frac{a}{x/s+1}\right)$$

using $z = x/s + 1$, one has

$$m_n = s^n\, e^b \int_1^{+\infty} dz\, (z-1)^n z^{-a} e^{-bz} \left(b + \frac{a}{z}\right)\ .$$

Taking the binomial expansion of the last expression and using (3) gives

$$m_n = \sum_{h=0}^{n} \binom{n}{h} s^n\, e^b\, \left(b\, \alpha_{h-a}(b) + a\, \alpha_{h-a-1}(b)\right)$$

which using (4) and the definition of the α function in terms of incomplete gamma, reduces to (7).

Q.E.D.

2.2 Inverse distribution function

The distribution function in (1) is defined over $[0, +\infty)$ and has image in $[0, 1)$. In this section we derive its inverse, the quantile function. This function, apart its theoretical interest, is useful for many purposes. For instance, it can be used to build so called *q-q* plots of *i.i.d.* samples or in in the computer generation of pseudo random number.

Consider the real function $f(x) = x\, e^x$. Since it is continuous and monotonically increasing for $x \geq -1$, it admits a continuous inverse $W(y)$ defined for $y \geq -1/e$. This function corresponds to the real branch of the Lambert function (c.f. Jeffrey *et al.* (1990)) and satisfy the equation

$$W(y)\, e^{W(y)} = y\ . \tag{8}$$

Theorem 2.3 *The inverse distribution function Q of* (1) *reads*

$$Q(q) = s\left(\frac{a}{b} W\left(\frac{b}{a}\, e^{\frac{b}{a}}\, (1-q)^{-1/a}\right) - 1\right)\ , \tag{9}$$

where $W(x)$ is the real branch of the Lamber function.

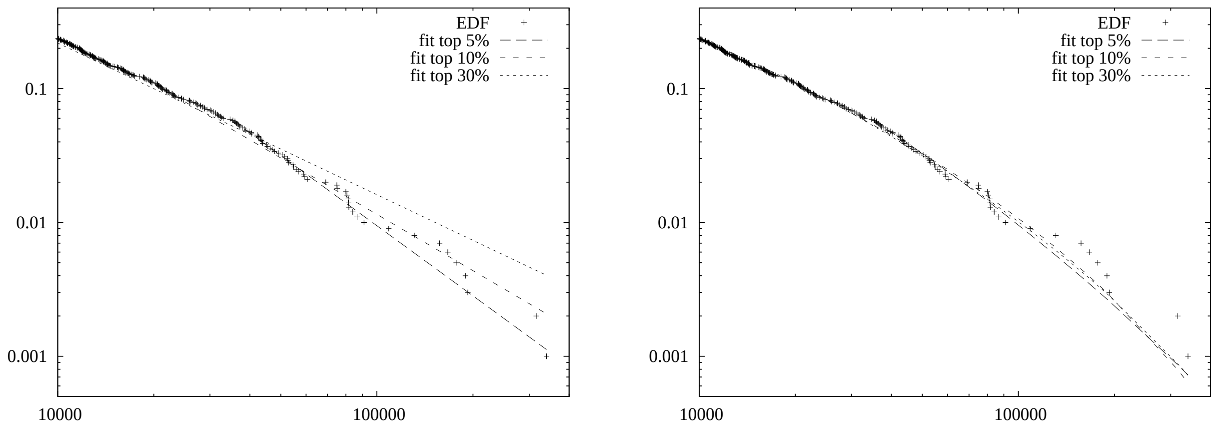

Figure 2: Distribution of the largest worldwide firms in 2006 according to Fortune 500. The maximum likelihood estimation of the upper tail obtained using 5%,10% and 30% of the available data is plotted for the Pareto Type I (left) and Pareto Type III (right) distribution.

	Pareto I			Pareto III		
sample size	5%	10%	30%	5%	10%	30%
$\hat{a}$	1.74	1.39	1.13	1.92	1.30	1.49
$\hat{s}/10^4$	0.68	0.41	0.26	1.07	0.49	0.63
$\hat{b}$	-	-	-	0.016	0.026	0.023

Table 1: Estimated coefficients for the Pareto Type I and III

Notice that when $q \in [0, +\infty)$, the argument of the function W in (9) is positive, so that the above expression is well defined. Moreover, for $q = 0$ the argument reduces to $e^{b/a}\, b/a$. From the definition of W, it is immediate to check that its value at this point is exactly a/b, so that one recover the relation $Q(0) = 0$.

Proof. Let $q \in [0, +\infty)$, we are interested in the value of x that satisfy $F(x) = q$. Considering the expression in (1) this reduces to the equation

$$\left(\frac{x}{s} + 1\right)^{-a} e^{-b\frac{x}{s}} = 1 - q \ .$$

Taking the logarithm of both sides and defining $z = b(x/s + 1)/a$, the previous expression reduces to

$$\log(x) + z = C \quad \text{where} \quad C = \frac{b - log(1-q)}{a} + \log(b) - \log(a) \ .$$

Taking the logarithm of (8), it is easy to see that the real solution of the previous equation is $z = W(\log C)$. Substituting the definition of C and remembering the previous change of variable, (9) follows.

Q.E.D.

As mentioned before, (9) can be straightforwardly used to generate *i.i.d.* pseudo-random variables extracted from a Pareto Type III distribution. To this purpose, one can generate a set of independent realizations $\{q_i\}$ uniformly distributed in $[0, 1)$ and apply the inverse distribution function Q to each realization to obtain a set $\{Q(q_i)\}$ of independent variates distributed according to (1).

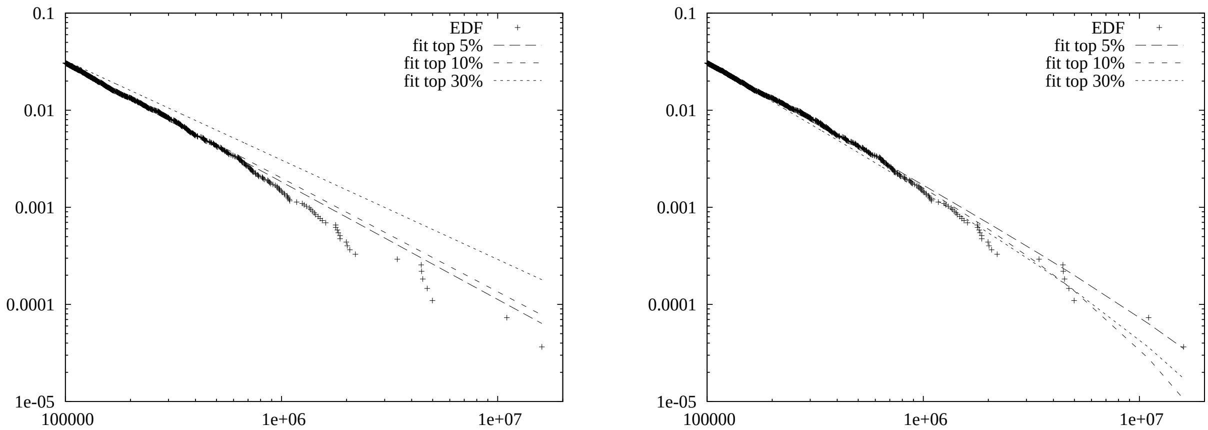

Figure 3: Distribution of the largest Italian manufacturing firms in 2004 (ISTAT data). The maximum likelihood estimation of the upper tail obtained using 5%,10% and 30% of the available data is plotted for the Pareto Type I (left) and Pareto Type III (right) distribution.

3 Firms size distribution

The upper tail exponential cut-off which characterizes the Pareto Type III distribution can prove extremely useful in obtaining better and more reliable descriptions of empirical data. We illustrate this claim with an example taken from Economics. In particular, we use the Pareto III distribution to describe the upper tail of the Empirical Distribution Function (EDF) of annual revenues of the largest worldwide companies. This choice seems particularly fit in our case, since it was exactly the problem of finding a reliable statistical description of the distribution of wealth which originally prompted the work of Pareto. We consider the Fortune 500 database, published each year by the Fortune magazine [1], which collects the revenues of the largest 500 firms in the world. Using maximum likelihood estimation (c.f. Hill (1975) and for details Bottazzi and Tamagni (2007)) we fit the Pareto Type III distribution defined in (1) to the EDF upper tail, using subpopulation of different sizes. For comparison, we also fit a Pareto Type I distribution defined as

$$F(x) = 1 - \left(\frac{x}{s}\right)^{-a} . \tag{10}$$

The Pareto Type I is the distribution traditionally applied to the description of the upper tail of wealth or income distribution (see for instance Castaldi and Milakovic (2007) and references therein). Results are reported in Fig. 2.2: the power-like decay of the Pareto I distribution shows up as a straight line in the log-log plot (left panel) while the exponential cut-off of the Pareto III distribution appears as a convex shape. As can be seen, the latter is remarkably less sensitive to the size of the considered sample and much more adapted to empirical observations. The estimate values for the tail index $\hat{a}$, the scale $\hat{s}$ and the exponential correction $\hat{b}$ are reported in Table 2.2. A similar analysis is repeated in Fig. 3 for the distribution of the largest Italian manufacturing firms using the database of firms with more than 20 employees provided by Italian Statistical office (ISTAT).

4 Conclusion

The Pareto Type III distribution, originally proposed by Pareto in 1896 as one of the possible statistical characterization of the distributional properties of the wealth in the Grand Duchy of Oldenburg, has never received much attention. However, its exponential asymptotic shape assures

[1]Data are publicly available at `http://money.cnn.com/magazines/fortune/fortune500/`

the existence of central moments and can be successfully used to probe for the presence of "deviation" from the power-like behavior of the upper tail of empirical distributions. This paper presents some properties of the Pareto III random variables. In particular, the formal expression of the inverse distribution function is derived, which allows for the simple computer generation of pseudo-random Pareto III variables.

References

Abramowitz, M. and I.A. Stegun (1964) *Handbook of Mathematical Functions*, New York: Dover.

Bottazzi, G. and Tamagni, F. (2007) "On the Uses and Misuses of Hill Estimator", LEM Working Papers, Scuola Superiore Sant'Anna, forthcoming.

Burroughs, S.M. and Tebbens, S.F. (2001), Upper-truncated power law distributions, *Fractals*, 9, 209-222.

Castaldi, C. and Milakovic, M. (2007), Turnover Activity in Wealth Portfolios, *Journal of Economic Behavior and Organization*, 63(3), 537-552.

Creedy, J. (1977), Pareto and the distribution of income, *Review of Income and Wealth*, 23, 405-411.

Hill, B.M. (1975), A Simple General Approach to Inference About the Tail of a Distribution, *The Annals of Statistics*, **3**, 1163-1174.

Kleiber C. and S. Kotz (2003), *Statistical Size Distributions in Economics and Actuarial Sciences.* Wiley and Sons Statistics: New York.

Newman M.E.J. (2005), Power laws, Pareto distributions and Zipf's law, *Contemporary Physics* 46, 323-351.

Pareto, V. (1986), Sur la Courve de la rèpartition de la richesse, Universitè de Lousanne. English traslation: *Revista di Politica Economica*, 87 (1997), 647-700.

Pareto, V. (1987), *Cours d'économie politique*, Lousanne: Ed. Rouge.

Jeffrey, D.J., Hare, D.E.G. and Corless, R. M. (1996), Unwinding the Branches of the Lambert W Function, *Math. Scientist* 21, 1-7.

Zipf, G.K. (1949), *Human Behaviour and the Principle of Least-Effort*, Cambridge MA: Addison-Wesley, 1949.