ADVANCES IN APPLIED PROBABILITY

VOLUME 17 NUMBER 2 JUNE 1985

APT

EDITOR-IN-CHIEF J. GANI

Adv. Appl. Prob. **17**, 465–469 (1985)
Printed in N. Ireland

AGGREGATION OF INDEPENDENT PARETIAN RANDOM VARIABLES

BERTRAND ROEHNER,* *Laboratoire de Physique Théorique et Hautes Energies,*† *Paris*
PETER WINIWARTER‡

Abstract

Empirical Paretian distributions play an important role in urban demography, size distributions of firms and income distributions; hence the addition of Paretian random variables is of interest. First, we give the asymptotic behavior (for large values of the variable) of the density function of a sum of n independently distributed Paretian variables. We then obtain the limiting distribution of an infinite sum of (i.i.d) Paretian variables and link our results with the theory of stable distributions.

1. Introduction

'Long-tailed' distribution functions are of interest in business, economics and the social sciences. Many processes were at first thought to be 'approximately normal' because tail observations were discarded as outliers. However, on retaining such observations one is often led to 'long-tailed' behavior very different from that of the normal law (Press (1982), p. 150).

The Paretian distribution is an example of such slowly decreasing distribution functions. Its density function is

$$f(x)=\alpha\frac{x_0^\alpha}{x^{\alpha+1}}Y(x-x_0)\qquad \alpha>0$$

where $Y(x)$ is given by

$$Y(x)=\begin{cases}0, & x<0\\ 1, & x\geqq 0.\end{cases}$$

The values of its moments $E(x)=m_1$, $D^2(x)=\sigma^2$ for different values of α are summarized below:

$$\begin{array}{lll} 0<\alpha\leqq 1 & m_1=\infty & \\ 1<\alpha\leqq 2 & m_1=\dfrac{\alpha}{\alpha-1}x_0, & \sigma^2=\infty \\ 2<\alpha & m_1=\dfrac{\alpha}{\alpha-1}x_0, & \sigma^2=\dfrac{\alpha x_0^2}{(\alpha-2)(\alpha-1)}. \end{array}$$

Received 9 July 1984; revision received 20 December 1984.
* Postal address: Université Paris VII, Tour 14–24, 5ème étàge, 2, place Jussieu, 75251 Paris Cedex 05, France.

† Laboratoire associé au CNRS LA 280.

‡ Postal address: Le Bordalier, F41270 Droué, France.

The Paretian curve is a good fit for the distribution of city sizes (Pumain (1982), Quandt (1964), Roehner and Wiese (1982), Zipf (1948)), of firm sizes (Hart and Prais (1956), Simon and Bonini (1958), Steindl (1965)) and of personal incomes (Mandelbrot (1960)), at least above a given threshold.

This would justify learning more about the behavior of a sum of Paretian variables; when the variance does not exist, however, for $1<\alpha\leqq 2$, the normal approximation does not apply. The mathematical problem we are addressing may then be stated as follows.

If $X_1, X_2, \cdots, X_n$ are independent but *non-identical* random variables with a Paretian distribution, what is the distribution of the sum $S(n)=\sum_{i=1}^{n} X_i$ and what happens if n goes to infinity?

We shall study the first part of the question in Section 2, or more precisely we shall give the asymptotic form of the density function of $S(n)$ for large values of the variable. This should be sufficient in practice, since Paretian distributions are always used above a given threshold.

The second question, closely related to the theory of stable distributions, will be answered in Section 3 for the case of i.i.d. Paretian variables.

2. Asymptotic behavior of a finite sum of Paretian variables

Proposition 1. Consider the sum $S(n)=X_1+\cdots+X_n$ of n independent *non-identical* Paretian random variables $\{X_i\}_{i=1}^{n}$ with density functions

$$f_{X_i}^{(x)}=\alpha_i\frac{(x_0^i)^{\alpha_i}}{x^{\alpha_i+1}}\,Y(x-x_0^i)\qquad \alpha_i>0,\qquad x_0^i>0.$$

Then the density function of the sum $S(n)$ has the following asymptotic behavior as $s\to\infty$:

$$\frac{f(s)}{S(n)}=\alpha_m\frac{(x_0^m)^{\alpha_m}}{s^{\alpha_m+1}}\qquad \alpha_m=\min(\alpha_1,\cdots,\alpha_n).$$

Proof. Since $f_{X_i}(x)$ is zero below x_0^i, we find for the Laplace transform $\varphi_{X_i}(\lambda)$ of $f_{X_i}(x)$:

$$\begin{aligned}\varphi_{X_i}(\lambda)&=\alpha_i(x_0^i)^{\alpha_i}\int_{x_0^i}^{\infty}\frac{\exp(-\lambda x)}{x^{\alpha_i+1}}\,dx\qquad \mathrm{Re}(\lambda)>0\\&=\alpha_i(x_0^i\lambda)^{\alpha_i}\Gamma(-\alpha_i,x_0^i\lambda)\end{aligned}$$

where the integral reduces to the incomplete gamma function (Magnus et al. (1966), p. 337).

We now expand the incomplete gamma function (Lösch (1951), p. 101):

$$\Gamma(a,x)=\Gamma(a)-x^a\sum_{n=0}^{\infty}\frac{(-)^n}{n!}\frac{x^n}{a+n}$$

which is valid for all real or complex a and x, except when a is a negative integer.

Hence we obtain

$$\varphi_{X_i}(\lambda)=1+\alpha_i\Gamma(-\alpha_i)(x_0^i\lambda)^{\alpha_i}-\frac{\alpha_i}{\alpha_i-1}(x_0^i\lambda)+\frac{\alpha_i}{2(\alpha_i-2)}(x_0^i\lambda)^2+\cdots.$$

The transform of the sum $S(n)$ is clearly

$$\varphi_{S(n)}(\lambda)=\prod_{i=1}^{n}\varphi_{X_i}(\lambda).$$

TABLE 1

Laplace transform	inverse
$\varphi_{S(n)}(\lambda) \underset{\lambda \to \lambda_0}{\simeq} (\lambda-\lambda_0)^\beta$; β real	$f_{S(n)}(s) \sim \dfrac{\exp(\lambda_0 t)}{\Gamma(-\beta)} \dfrac{1}{t^{\beta+1}}$; $\dfrac{1}{\Gamma(-\beta)} = 0$ for $\beta = 0, 1, 2$
$\varphi_{S(n)}(\lambda) \underset{\lambda \to \lambda_0}{\simeq} (\lambda-\lambda_0)^\beta \ln(\lambda-\lambda_0)$; $\beta \neq 0, 1, 2, \cdots$	$f_{S(n)}(s) \sim \left(\dfrac{d}{d\beta} \dfrac{1}{\Gamma(-\beta)}\right) \exp(\lambda_0 t) \dfrac{1}{t^{\beta+1}}$
$\varphi_{S(n)}(\lambda) \underset{\lambda \to \lambda_0}{\simeq} (\lambda-\lambda_0)^p \ln(\lambda-\lambda_0)$; $p = 0, 1, 2, \cdots$	$f_{S(n)}(s) \sim (-)^{p+1} p! \exp(\lambda_0 t) \dfrac{1}{t^{p+1}}$

We now use a result of Doetsch ((1950), pp. 498, 501) giving the asymptotic behavior of the inverse of $\varphi_{S(n)}(\lambda)$, that is the density function $f_{S(n)}(s)$ of the sum $S(n)$.

If $\varphi_{S(n)}(\lambda)$ has only one singular point λ_0, then the asymptotic behavior, as $s \to \infty$, of its inverse is given by $f_{S(n)}(s)$ as shown in Table 1.

When none of the α_i is an integer, we shall need only the first part of the previous result. It shows that the asymptotic behavior of $f_{S(n)}(s)$ is given by the lowest non-integral power in the expansion of $\varphi_{S(n)}(\lambda)$. Since $\lambda = 0$ is the only singular point of the $\varphi_{X_i}^{*}(\lambda)$, we find as $s \to \infty$:

$$f_{S(n)}(s) \sim \alpha_m \frac{(x_0^m)^{\alpha_m}}{s^{\alpha_m+1}}$$

where $i = m$ denotes the index of the smallest α_i. When α_i is an integer, it can be shown that, apart from a constant factor, the behavior of $f_{S(n)}(s)$ is the same.

3. Behavior of an infinite sum of i.i.d. Paretian variables: A central limit theorem

Let us now see what happens when $n \to \infty$. More precisely, we investigate the limit of the density function of a 'renormalized' sum:

$$S_1(n) = \frac{X_1 + \cdots + X_n}{a_n} - b_n$$

where a_n and b_n are adjusted to ensure that the limit of $S_1(n)$ exists. We shall restrict ourselves to i.i.d. Pareto variables. The problem is clearly related to the theory of stable random variables (Gnedenko and Kolmogorov (1954)). The family of stable distributions coincides with the set of all distributions which are a limit of sums like $S_1(n)$ of independent and identically distributed variables.

Restricting the X_i to Paretian variables results in a special stable distribution. The general expression for the characteristic function of all stable distributions is well known; however, the canonical representation does not tell us how to choose the right a_n and b_n, nor does it give us the relation between the distribution function of the variables X_i and the distribution of their limit. In this section we prove the following central limit theorem.

Proposition 2. The density function of the sum

$$S_1(n) = \frac{X_1 + \cdots + X_n}{a_n} - b_n$$

where the X_i are identically and independently distributed Pareto variables with power

α, converges to the function

$$W(s)=\frac{1}{\alpha\pi}\int_0^{\infty}\exp(-au)\cos(u^{1/\alpha}s+bu)u^{(1/\alpha)-1}\,du$$

where $a=-x_0^{\alpha}\alpha\Gamma(-\alpha)\cos(\alpha\pi/2)$, $b=x_0^{\alpha}\alpha\Gamma(-\alpha)\sin(\alpha\pi/2)$ provided the coefficients a_n, b_n are chosen so that

$$\begin{array}{lll} \text{if } 1<\alpha<2, & a_n=n^{1/\alpha}, & b_n=nE(X_1) \\ \text{if } 0<\alpha<1, & a_n=n^{1/\alpha}, & b_n=0. \end{array}$$

Proof.

Case $1<\alpha<2$. Since the mean m_1 exists in this case, we replace the variables X_i by the corresponding centred variables:

$$\hat{X}_i=\frac{X_i-m_1}{a_n}$$

where the renormalization coefficients a_n are unknown for the moment.

Since the $\hat{X}_i$ are no longer restricted to positive values, we now use the characteristic function:

$$\psi_{X_i}(t)=\int_{-\infty}^{+\infty}\exp(itx)f_{X_i}(x)\,dx=\varphi_{X_i}(-it).$$

We get thus:

$$\psi_{\hat{X}_i}=\exp\left(-\frac{itm_1}{a_n}\right)\varphi_{X_i}\left(-\frac{it}{a_n}\right).$$

Expanding the exponential and the incomplete gamma functions about $t=0$ yields

$$\psi_{\hat{X}_i}(t)=\left[1-\frac{itm_1}{a_n}-\frac{1}{2}\left(\frac{tm_1}{a_n}\right)^2+\cdots\right]\left[1+\frac{\alpha}{\alpha-1}\frac{itx_0}{a_n}+\alpha\Gamma(-\alpha)\left(\frac{-itx_0}{a_n}\right)^{\alpha}+\cdots\right]$$

it is easy to see that with $a_n=n^{1/\alpha}$ we obtain the following limit:

$$\begin{aligned}\psi_{S_j(n)}(t)=\lim_{n\to\infty}\psi^n_{\hat{X}_i}(t)&=\exp\left[\alpha\Gamma(-\alpha)\left(\frac{tx_0}{i}\right)^{\alpha}\right]\\ &=\exp\left[\alpha x_0^{\alpha}\Gamma(-\alpha)\cos\alpha\frac{\pi}{2}|t|^{\alpha}\left(1-i\operatorname{sgn}t\,.\,\operatorname{tg}\alpha\frac{\pi}{2}\right)\right].\end{aligned}$$

If we compare this result with the general expression of the characteristic function of a stable distribution given in Gnedenko and Kolmogorov (1954), p. 164:

$$\psi(t)=\exp\left[i\gamma t-c\,|t|^{\alpha}\left(1+i\beta\operatorname{sgn}t\,.\,\operatorname{tg}\alpha\frac{\pi}{2}\right)\right]$$

we see that

$$\gamma=0;\qquad \beta=-1;\qquad c=-\alpha\Gamma(-\alpha)x_0^{\alpha}\cos\alpha\frac{\pi}{2}>0.$$

Going now to the inverse $W(s)$ of the Fourier transform readily leads to the expression given in Proposition 2.

Case $0<\alpha<1$. The mean m_1 no longer exists, so we simply renormalize the X_i to

$\tilde{X}_i = X_i/a_n$, where

$$\psi_{\tilde{X}_i}(t) = 1 + \alpha\Gamma(-\alpha)\left(-\frac{itx_0}{a_n}\right)^{\alpha} + \frac{\alpha}{\alpha-1}\frac{itx_0}{a_n} + \cdots.$$

The same choice of a_n works.

Acknowledgement

The authors would like to express their gratitude to Drs G. Valent and K. E. Wiese for their kind assistance.

References

Doetsch, G. (1950) *Handbuch der Laplace-Transformation*, Vol. 1. Birkhaüser, Basel.

Gnedenko, B. V. and Kolmogorov, A. N. (1954) *Limit Distributions for Sums of Independent Random Variables*. Addison-Wesley, Reading, Mass.

Hart, P. E. and Prais, S. J. (1956) The analysis of business concentration: a statistical approach. *J. R. Statist. Soc.* A**119,** 150–191.

Lösch, F. and Schoblick, F. (1951) *Die Fakultät* (*Gamma Funktion*) *und verwandte Funktionen*. Teubner, Leipzig.

Magnus, W., Oberhettinger, F. and Soni, R. P. (1966) *Formulas and Theorems for the Special Functions of Mathematical Physics*. Springer-Verlag, New York.

Mandelbrot, B. (1960) The Pareto–Lévy law and the distribution of income. *Internat. Econom. Rev.* **1,** 79–106.

Press, S. J. (1982) *Applied Multivariate Analysis*. R. E. Krieger, Malabar, Florida.

Pumain, D. (1982) *La dynamique des villes*. Economica, Paris.

Quandt, R. E. (1964) Statistical discrimination among alternative hypothesis and some economic regularities. *J. Regional Sci.* **5**(2), 1–23.

Roehner, B. and Wiese, K. E. (1982) A dynamic generalization of Zipf's rank-size rule. *Environment and Planning* A **14,** 1449–1467.

Simon, H. A. and Bonini, C. P. (1958) The size distribution of business firms. *American Econom. Rev.* **48,** 607–617.

Steindl, J. (1965) *Random Processes and the Growth of Firms*. Hafner, New York.

Zipf, G. K. (1948) *Human Behavior and the Principle of Least Effort*. Addison-Wesley, Cambridge, Mass.