The Annals of Statistics
1974, Vol. 2, No. 3, 599–601

A CHARACTERIZATION OF THE PARETO DISTRIBUTION

BY NAGESH S. REVANKAR, MICHAEL J. HARTLEY
AND MARCELLO PAGANO

State University of New York at Buffalo

A positive random variable X whose mean exists, has a Pareto distribution if, and only if, $E(X \mid X > x) = h + gx$ for $g > 1$. This characterization was motivated by the fact that the Pareto distribution has been widely used to model income. Now, suppose that individuals under-report their income for income tax purposes. If one assumes that for a given income, the average under-reporting is a constant fraction of the amount by which the income exceeds the tax exempt level, then the average under-reporting error for a given reported income is a linear function of the reported income if, and only if, incomes follow a Pareto distribution.

1. Motivation. The determination of the laws of income distribution has long occupied a prominent place in economics. One of these—the Pareto Law—asserts that:

> The logarithm of the percentage of units with an income exceeding some value is a negatively sloped linear function of the logarithm of that value.

The resulting Pareto distribution is usually assumed to adequately represent the distribution of incomes above certain "low" values. As Klein (1962, page 151) suggests we may consider the law as applicable to taxpayers whose incomes exceed tax-exempt levels. Those incomes below such a level, say m, are excluded from characterization by the Pareto Law.

It is also common that individuals may under-report their true incomes to avoid payment of some portion of their income tax. Clearly, when incomes fall below the tax exempt level m, there is no "incentive" to under-report. However, for incomes $X > m$, evidence indicates that under-reporting, though illegal, is not unusual. It is, therefore, of some interest to be able to characterize the Pareto Law in the presence of this under-reporting "error". Needless to say, such a characterization will involve some assumptions regarding the aforementioned error.

Let the random variables X, Y and U denote the actual income, reported income and the under-reporting error, respectively. We define

$$Y = X - U \quad \text{where} \quad 0 < U < \max(0, X - m) . \tag{1.1}$$

This restriction is equivalent to assuming $Y > m$ whenever the associated $X > m$. It appears to be a reasonable assumption provided the lawbreaker is "rational". Without outlining the underlying behavioral model, presumably

Received April 1972; revised July 1973.
AMS 1970 *subject classifications*. Primary 62E10; Secondary 62P20.
Key words and phrases. Pareto distribution, characterization theorem, linear regression.

rational under-reporting involves maximizing the difference between expected benefits (savings in taxes due to a lower declaration) and expected costs (the probability of detection multiplied by the penalty, if convicted). If the expected costs vary directly with the magnitude of the reporting error, then, since no additional benefits are obtained by reporting $Y < m$ when $X > m$ (whereas additional costs may be expected), we would expect to observe an associated Y also greater than m.

We assume that the average amount of under-reporting from a given $X = x > m$, is proportional to $x - m$, i.e.,

$$E(U \mid X = x) = b(x - m) = a + bx \tag{1.2}$$

where $0 < b < 1$ and $a = -bm$.

An immediate consequence of (1.1) is that when X has Pareto distribution on (m, ∞), with probability one, Y takes on values in (m, ∞).

The Pareto Law has the distribution function $F(\cdot) = 1 - G(\cdot)$, where

$$G(x) = ((m + c)/(x + c))^{\theta} \qquad x > m\,, \tag{1.3}$$

one elsewhere, where $\theta > 0$ and the location parameter, $c > -m$.

In order that the Pareto Law have a finite mean, θ must be greater than one. Since our characterization, in Section 2, is in terms of a regression function, we shall make this assumption ($\theta > 1$) henceforth. This restriction does not seem to hinder the applicability of the Pareto Law since there is general agreement from empirical studies that $\theta > 1$ in practice; see for example Klein (1962) and Aigner and Heins (1967). In the latter study, separate analyses of income data for 50 states in 1960 yielded estimates of θ between 2.2 and 3.1.

The Pareto Law may be subdivided into two types depending on the value of the location parameter c. If $c = 0$, (1.3) is defined to be a Pareto Type 1 distribution. For $c > -m$, (1.3) defines a Pareto Type 2 distribution.

2. A characterization.

Consider observing a reported income of y. We have the

THEOREM. *Let* (1.1) *and* (1.2) *hold. Then for*

$$E(U \mid X > y) = \alpha + \beta y \tag{2.1}$$

with $\beta > b > 0$, *it is necessary and sufficient that* X *have a Pareto distribution with finite mean.*

PROOF. The sufficiency is straightforward. To prove the necessary part, we must show that the only distribution $F(\cdot)$ (or $G(\cdot) = 1 - F(\cdot)$) for which

$$\frac{1}{G(y)} \int_y^\infty (a + bx)\, dF(x) = (\alpha + \beta y) \qquad \forall y > m \tag{2.2}$$

with $\beta > b$, is that given by (1.3).

We first show that $F(\cdot)$ is necessarily continuous. An immediate inequality obtainable from (2.2) is that $a + by \leqq \alpha + \beta y$ for all $y > m$ and in particular

$$m \geqq (a - \alpha)/(\beta - b)\,. \tag{2.3}$$

Now suppose that $F(\cdot)$ is not continuous and it has a discrete part. Suppose $F(\cdot)$ has a saltus at y_0. Then it is readily seen from (2.2) that y_0 is unique and, furthermore, $y_0 = (a - \alpha)/(\beta - b)$. But from (2.3) this is outside our range of consideration, so $F(\cdot)$ is continuous.

Define the function

$$H(y) = \textstyle\int_y^\infty G(x)\,dx\,, \qquad y > m\,. \tag{2.4}$$

Since $F(\cdot)$ ($G(\cdot)$) is continuous, $H(\cdot)$ is differentiable and $H'(\cdot) = -G(\cdot)$. Integrating (2.2) by parts we see that

$$\frac{d \log H(y)}{dy} = \frac{-b}{\alpha - a + (\beta - b)y}\,. \tag{2.5}$$

Solving this differential equation yields the theorem.

An immediate corollary is:

COROLLARY. *If $a = \alpha$, then the theorem holds true for a Pareto Type 1 distribution.*

D. R. Cox (1962, page 128) has the result that the conditional expectation $E(X \mid X > x)$ for a positive random variable (when it exists) characterizes the distribution of the random variable. Our theorem considers a specific functional form for this conditional expectation. The theorem is, however, more general than the one established by Krishnaji (1970). He considers a specific distribution for the under-reporting error and provides a characterization of Pareto Type 1 distribution. Incidentally, if we let $\beta = b$, then, necessarily with $\alpha > a$ we have another functional form for the conditional expectation which yields a well-known characterization of the exponential (see, for example, Reinhardt (1968), or Shanbhag (1970)).

REFERENCES

[1] AIGNER, D. J. and HEINS, A. J. (1967). A social welfare view of the measurement of income equality. *Rev. Income Wealth Ser.* 13 **1** 12–26.

[2] COX, D. R. (1962). *Renewal Theory*. Methuen, London.

[3] KLEIN, L. R. (1962). *Introduction to Econometrics*. Prentice Hall, New York.

[4] KRISHNAJI, N. (1970). Characterization of the Pareto distribution through a model of under-reported incomes. *Econometrica* **38** 251–255.

[5] QUANDT, R. E. (1966). Old and new methods of estimation and the Pareto distribution. *Metrika* **10** 55–82.

[6] REINHARDT, H. E. (1968). Characterizing the exponential distribution. *Biometrics* **24** 437–439.

[7] SHANBHAG, D. N. (1970). Characterizations for exponential and geometric distributions. *J. Amer. Statist. Assoc.* **65** 1256–1259.

ECONOMICS DEPARTMENT AND
STATISTICAL SCIENCE DIVISION
STATE UNIVERSITY OF NEW YORK
AMHERST, NEW YORK 14226