Estimation of the Parameters of the Pareto Distribution[1])

By H. J. Malik, Guelph[2])

Summary: In this paper, sufficient statistics for the parameters a and v of the Pareto distribution are obtained. Using sufficiency, it is shown that the statistic

$$Z = \sum_{i=1}^{N} \log \left({}^{Y_i}/{Y_1} \right)$$

is stochastically independent of the sufficient statistic Y_1. Using sufficiency and stochastic independence of Z and Y_1, the exact distribution of the maximum likelihood estimator $\hat{v}$ is derived.

1. Introduction

It is well known that a Pareto distribution

$$f(x) = va^v x^{-v-1} \quad a > o,\ v > o,\ x \geq a \tag{1.1}$$

Provides reasonably good fits to distribution of income and of property values. For detailed arguments on the existence of such distributions in economic life the reader is referred to the discussions by Davis, Hagstroem [1925, 1960] and Mandlebrot [1963].

Let $X_1, \ldots, X_N$ be a random sample of size N from (1.1). Let $Y_1 < Y_2 < \ldots < Y_N$ represent $X_1, \ldots, X_N$ when the latter are arranged in ascending order of magnitude. Then Y_i $(i = 1, 2, \ldots, N)$ is called the ith order statistic of the random sample $X_1, \ldots, X_N$. Then it is easily seen [9] that the maximum likelihood estimate of a is $Y_1 = \text{Min}\,(X_1, \ldots, X_N)$ and that of v is

[1]) Paper presented at the Eastern Regional meeting of the Institute of Mathematical Statistics, Upton, Long Island, New York, April 27—29, 1966.

[2]) H. J. Malik, University of Guelph, Guelph, Ontario, Canada. Work done when the author was on the faculty of Western Feserve University, Cleveland, Ohio, USA.

$$\hat{v} = \left[\log \frac{g}{Y_1}\right]^{-1}$$

In this paper, sufficient statistics for the parameters a and v of the PARETO distribution are obtained. It is shown that $Y_1 = \text{Min}\,(X_1, \ldots, X_N)$ is sufficient for a when v is known; the sample geometric mean g is sufficient for v when a is known, and $\left(Y_1, \sum_{i=1}^{N} \log \frac{Y_i}{Y_1}\right)$ is a joint set of sufficient statistics for (a, v), when both are unknown, where $(Y_1, Y_2, \ldots, Y_N)$ are the order statistics of the random sample. Using sufficiency, it is shown that the statistic $Z = \sum_{i=1}^{N} \log \frac{Y_i}{Y_1}$ is stochastically independent of the sufficient statistic Y_1. Using sufficiency and stochastic independence of Z and Y_1, the exact distribution of the maximum likelihood estimator $\hat{v}$ is derived.

2. Sufficient Statistics and Stochastic independence

Theorem 1: Let $X_1, X_2, \ldots, X_N$ be a random sample of size N that has p. d. f.

$$f(x) = v a^v x^{-v-1} \qquad a > o,\ v > o,\ x \geq a \tag{2.1}$$

Then the sample geometric mean g is a sufficient statistic for v for known a and $Y_1 = \text{Min}\,(X_1, X_2, \ldots, X_N)$ is a sufficient statistic for a for known v and $\left(\sum_{i=1}^{N} \log \frac{Y_i}{Y_1}, Y_1\right)$ is a joint set of sufficient statistics for (v, a) when both are unknown.

Proof: (i) a *Known*

It can be easily shown that the statistic $U = \log X_1 + \log X_2 + \ldots + \log X_N$ has the p. d. f.

$$g\,(u; v) = \frac{v^N}{\Gamma\,(N)}\,(u - N \log a)^{N-1} e^{-v\,(u - N \log a)}$$

$$u - N \log a > o$$

Accordingly, the joint p. d. f. of $X_1, \ldots, X_N$ from (2.1) may be written as

$$v^N a^{Nv} x_1^{-v-1} \ldots x_N^{-v-1} = \left[\frac{v^N}{\Gamma\,(N)}\,(u - N \log a)^{N-1} e^{-v(u - N \log a)}\right] \left[\frac{\Gamma\,(N)}{(u - N \log a)^{N-1}\, x_1 \ldots x_N}\right]$$

$$= g(u; v)\, H(x_1, x_2, \ldots, x_N)$$

In accordance with the FISHER-NEYMAN criterion [HOG, CRAIG, p. 262], $U = \log X_1 + \log X_2 + \ldots + \log X_N$ is a sufficient statistic for v for known a. Since the sample geometric mean g is a function of u and N only, the sample geometric mean g is also a sufficient statistic for known a.

(ii) v *Known*

The statistic $Y_1 = \text{Min}\ (X_1, X_2, \ldots, X_N)$ i. e., the first order statistic in a random sample of size N has the p. d. f.

$$g(y_1) = Nva^{Nv}\, y_1^{-Nv-1} \qquad a < y_1 < \infty$$

The joint p. d. f. of $X_1, X_2, \ldots, X_N$ from (2.1) may be written

$$v^N a^{Nv} x_1^{-v-1} \ldots x_N^{-v-1} = \left[Nva^{Nv} y_1^{-Nv-1}\right]\left[\frac{1}{N} v^{N-1} (x_1 x_2 \ldots x_N)^{-v-1} y^{Nv+1}\right]$$
$$= g(y_1, a)\, H(x_1, x_2, \ldots, x_N).$$

In accordance with the FISHER-NEYMAN criterion [HOGG, CRAIG, p. 262], $Y_1 = \text{Min}\ (X_1, X_2, \ldots, X_N)$ is a sufficient statistic for a for known v.

(iii) *Both unknown*. The p. d. f. of $U = \sum_{i=1}^{N} \log \frac{Y_i}{Y_1}$ derived in Theorem 3, is

$$f(z) = \frac{v^{N-1}}{\Gamma(N-1)} z^{N-2} e^{-vz} \qquad z > o$$

On the other hand, the likelihood of the sample may be rewritten as

$$L = v^N a^{Nv} y_1^{-v-1} \ldots y_N^{-v-1}$$
$$= \left[\frac{v^{N-1}}{\Gamma(N-1)} z^{N-2} e^{-vz}\right]\left[Nva^{Nv} y_1^{-Nv-1}\right] H(y_1, \ldots, y_N)$$
$$= g(z, y_1; a, v)\, H(y_1, \ldots, y_N)$$
$$= g(z, y_1; a, v)\, H(x_1, \ldots, x_N).$$

In accordance with the FISHER-NEYMAN criterion [HOGG, CRAIG, p. 262] $\left(Z = \sum_{i=1}^{N} \log \frac{Y_i}{Y_1},\ Y_1\right)$. is a joint set of sufficient statistics for (v, a).

Theorem 2: Let $X_1, X_2, \ldots, X_N$ denote a random sample of size N from a distribution having a p. d. f.

$$f(x) = va^v x^{-v-1} \qquad a > o,\ v > o,\ x \geq a$$

Let $Y_1 < Y_2 < \ldots < Y_N$ denote the order statistics of this random sample. Then $Z = \sum_{i=1}^{N} \log \frac{Y_i}{Y_1}$ is stochastically independent of the sufficient statistic Y_1.

Proof:

Since the distribution of Z is uniquely determined by its moment generating function $M_Z(t)$. To establish the independence of Z and Y_1, we have to show that $M_Z(t)$ does not depend upon a, since the p. d. f. of Y_1 is complete.

The moment generating function $M_Z(t)$ is given by

$$M_Z(t) = \int_a^\infty \int_a^{y_n} \dots \int_a^{y_2} e^{t \sum_{i=1}^{N} \log \frac{Y_i}{Y_1}} N!\, v^N a^{Nv} \prod_{i=1}^{N} y_1^{-v-1} dy_i$$

Make the transformation

$$w_1 = \frac{a}{y_1},\ w_2 = \frac{a}{y_2},\ \dots,\ w = \frac{a}{y_N}.$$

The transformation is one-to one, $|J| = a^N$, and it maps

$$a < y_1 < y_2 < \dots < y_N < \infty$$

into

$$o < w_N < w_{N-1} < \dots < w_1 < 1$$

consequently,

$$M_Z(t) = \int_0^1 \int_0^{w_1} \dots \int_0^{w_{N-2}} \int_0^{w_{N-1}} e^{t \sum_{i=1}^{N} \log \frac{w_1}{w_i}} N! a^{Nv} v^N \prod_{i=1}^{N} \left(\frac{a}{w_i}\right)^{-v-1} dw_i \mid J \qquad (2.2)$$

Substituting $|J| = a^N$ in (2.2), we get

$$M_Z(t) = \int_0^1 \int_0^{w_1} \dots \int_0^{w_{N-2}} \int_0^{w_{N-1}} e^{t \sum_{i=1}^{N} \log \frac{w_1}{w_i}} N! v^N \prod_{i=1}^{N} w_i^{v+1}\, dw_i$$

which clearly does not depend upon a. Thus, by a Theorem given in [HOGG, CRAIG p. 232], the distribution of Z does not depend upon a, and so

$$Z = \sum_{i=1}^{N} \log \frac{Y_i}{Y_1}$$

is stochastically independent of Y_1, the sufficient statistic for a, since the p. d. f. of Y_1 is complete.

3. Distribution of the Maximum Likelihood Estimator of v

Theorem 3: Let $X_1, X_2, \dots, X_N$ be a random of size N from a distribution having a p. d. f.

$$f(x) = v a^v x^{-v-1} \qquad a > o,\ v > o,\ x \geq a.$$

Let g be the sample geometric mean and $Y_1 = \text{Min}\,(X_1, \dots, X_N)$. Then the p. d. f. of $S = \log \frac{g}{Y_1}$ is given by

$$g(s) = \frac{v^{N-1}\, N^{N-1}}{\Gamma(N-1)}\, s^{N-2}\, e^{-vNs} \qquad s > o \tag{3.1}$$

Proof:

We make the observation that

$$\sum_{i=1}^{N} \log \frac{X_i}{X_1} = \log X_2 + \ldots + \log X_N - (N-1) \log X_1$$

does not depend on the ordering of $X_2, X_3, \ldots, X_N$. Thus, if we take $X_1 < X_2, X_3, \ldots, X_N$

$$\sum_{i=1}^{N} \log \frac{Y_i}{Y_1} = \sum_{i=1}^{N} \log \frac{X_i}{X_1}$$

and the conditional distribution of $X_2, X_3, \ldots, X^N$, given $X_1 = x_1$, is

$$g(x_2, \ldots, x_N \mid x_1) = \frac{f(x_2) \ldots f(x_N)}{[1 - F(x_1)]^{N-1}} \qquad x_1 < x_2, \ldots, x_N$$

The characteristic function of

$$Z = \sum_{i=1}^{N} \log \frac{Y_i}{Y_1} = \sum_{i=1}^{N} \log \frac{X_i}{X_1} \text{ given } X_1 = x_1 \text{ is}$$

$$\emptyset(t) = E\left[e^{it \Sigma \log \frac{X_i}{X_1}} \,\Big|\, x_1 \right] = \left[\int_{x_1}^{\infty} \frac{e^{it \log \frac{x_2}{x_1}} f(x_2)\, dx_2}{1 - F(x_1)} \right]^{N-1}$$

This follows from the fact that each X_i, $i = 2, 3, \ldots, N$, given $X_1 = x_1$, has the same distribution and the X_i, $i = 2, 3, \ldots, N$ are conditionally mutually independent.

The p. d. f. of Z is given by

$$f(z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itz}\, \emptyset(t)\, dt$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itz} \left[\int_{x_1}^{\infty} \frac{e^{it \log \frac{x_2}{x_1}} f(x_2)\, dx_2}{1 - F(x_1)} \right]^{N-1} dt \tag{3.2}$$

Substituting $F(x_1) = 1 - a^v x_1^{-x}$, and after some algebraic simplifications (3.2) reduces to

$$f(u) = \frac{v^{N-1}}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-itz}\, dt}{(v - it)^{N-1}} \tag{3.3}$$

It is shown in [8] that

$$\int_{-\infty}^{\infty} \frac{e^{-ibx}\,dx}{(r-ix)^a} = \frac{2\,\pi}{\Gamma(a)}\, b^{a-1}\, e^{-vz} \tag{3.4}$$

Now using (3.4), we have

$$f(z) = \frac{v}{\Gamma(N-1)}\, z^{N-2}\, e^{-vz} \qquad z > o \tag{3.5}$$

(3.2) is the conditional distributions of $Z = \sum_{i=1}^{N} \log \frac{Y_i}{Y_1}$ given Y_1. Since Y_1 is a sufficient statistic and since the p. d. f. of Y_1 is complete. By Theorem 2, $Z = \sum_{i=1}^{N} \log \frac{Y_i}{Y_1}$ and Y_1 are independent and thus the conditional distribution of $Z = \sum_{i=1}^{N} \log \frac{Y_i}{Y_1}$ given $Y_1 = y_1$ is equal to the unconditional distribution. So the p. d. f. of $Z = \sum_{i=1}^{N} \log \frac{Y_i}{Y_1}$ is given by (3.5). But $Z = \sum_{i=1}^{N} \log \frac{Y_i}{Y_1}$ can be rewritten as

$$Z = \sum_{i=1}^{N} \log \frac{Y_i}{Y_1} = \log Y_1 + \log Y_2 + \ldots + \log Y^N - N \log Y_1$$

$$= \log \frac{Y_1\, Y_2 \ldots Y_N}{Y_1^N} = \log \left[\frac{(Y_1 Y_2 \ldots Y_N)^{\frac{1}{N}}}{Y_1}\right]^N$$

or

$$Z = N \log \frac{g}{Y_1}.$$

Setting $\log \frac{g}{Y_1} = S$ and making the transformation $Z = NS$, (3.5) reduces to

$$g(s) = \frac{v^{N-1}\, N^{N-1}}{\Gamma(N-1)}\, s^{N-2}\, e^{-vNs} \qquad s > o.$$

Corollary: The p. d. f. of the maximum likelihood estimator $\hat{v} = \left[\log \frac{g}{Y_1}\right]^{-1}$ is given by

$$g(\hat{v}) = \frac{v^{N-1}\, N^{N-1}\, 1}{\Gamma(N-1)\, (v)^N}\, e^{-\frac{v}{v} N} \qquad \hat{v} > o \tag{3.6}$$

Proof:

Making the transformation $\hat{v} = \frac{1}{s}$ in (3.1) and multiplying by the

Jacobian, $|J| = \frac{1}{\hat{v}}$, we get (3.6) the p. d. f. of the maximum likelihood estimator $\hat{v}$. We also note that the distribution of $\frac{2Nv}{\hat{v}}$ is x^2-distributed with $2N-2$ degrees of freedom.

References

DAVIS, H. T.: The analysis of economic time series, The Cowles Commission for Research in Economics, Monograph No. **6**, Bloomington.

HAGSTROEM, K. G.: Remarks on Pareto Distributions, Skandinavisk Aktuardits-skrift, 1960, p. 59–71.

— L a loi de Pareto et la reassurance, Skandinavisk Aktuartidskrift, 1925, p. 65.

HOOG, R. V., A. T. CRAIG: Introduction to Mathematical Statistics, New York, 1964.

MALIK, H. J.: Exact moments of order statistics from the Pareto distribution, (ba-stract) Ann. Math. Statist. **37**, 1966, p. 553.

— Estimation of the Parameters of the Pareto distribution. (abstract) Ann. Math. Statist. **37**, 1966, p. 765.

MANDLEBROT, B.: New methods in Statistical Economics, Bull. Inst. Int. Stat., 1963, XL Book 2.

MACROBERT, T. M.: Functions of a Complex Variable, London 1967.

QUANDT, R. E.: Old and new methods of estimation and the Pareto Distribution, Metrika, **10**, 1966, p. 55–82.