Ann. Inst. Statist. Math.
28 (1976), Part A, 491-497

CHARACTERIZING THE PARETO AND POWER DISTRIBUTIONS

A. C. DALLAS

(Received Apr. 9, 1973)

1. Introduction

Let X be a random variable with distribution function

$$F(x)=\left(\frac{x}{a}\right)^{\theta}, \qquad 0\leqq x\leqq a \text{ and } \theta>0 . \tag{1.1}$$

Then X has the power distribution. If Y has as distribution function

$$G(y)=1-\left(\frac{\beta}{y}\right)^{\gamma}, \qquad 0<\beta\leqq y \text{ and } \gamma>0 \tag{1.2}$$

then Y has the Pareto distribution. It can shown that if X has the power distribution with parameters a and θ then X^{-1} has the Pareto distribution parameters $\beta=a^{-1}$ and $\gamma=\theta$. These two distributions are also related with the exponential distribution, on account of the following transformations. If Z has the exponential distribution

$$\Gamma(z)=1-e^{-\lambda z}, \qquad z\geqq 0,\ \lambda>0 \tag{1.3}$$

then (*) $Y=\beta e^{Z}$ has the Pareto and (**) $X=ae^{-Z}$ has the power distribution. These two transformations are one to one.

Many characterizations of the Pareto and power distributions are known. We observe that characterizations of the exponential distribution based on distributional properties, can be transferred to the power and Pareto distributions due to the previously mentioned transformations of the exponential variable. The power distribution is characterized in [1], [2], [3] and the Pareto distribution in [4]. In these papers, characterizations for the exponential distribution are given, using independence properties of the order statistics. Then using (*) and (**) the desired characterizations follow for the power and Pareto distributions. The same also holds for the lack of memory property of the exponential distribution. This property can be stated as follows: "If Z has (1.3) as distribution, the conditional distribution of Z given that $Z>c$ is the same as the distribution of $Z+c$ for all $c>0$." By (*) we conclude for the power distribution that X, given $X\leqq c$, has the same distribution as cX/a all c, $0<c<a$. Similarly, Y given that $Y>c$ and Yc/β are identically distributed for all $c>\beta$, if Y has the Pareto distribution. As the transformations are one to one and as the lack of memory characterizes (1.3), we have that the previous properties characterize (1.1) and (1.2).

A direct characterization for the power distribution is given in [5] by M. Fisz. He considers independence properties of order statistics. In [6], [7] and [8], independence is replaced by the weaker property of constant conditional expectation. Krishnaji N. [9] characterizes the Pareto distribution through the distribution of incomes and Srivastava M. S. [10] characterizes the power and Pareto distributions using the mean and the extreme observations of the sample.

2. Summary

In this note we prove that

$$\text{(2.1)} \qquad \mathrm{E}\,(X^r|X\leqq c)=\mathrm{E}\left(\frac{Xc}{a}\right)^r \qquad 0\leqq x\leqq a \text{ all } c \text{ in } (0,a),\ r>0$$

characterizes the power distribution. Similarly,

$$\text{(2.2)} \qquad \mathrm{E}\,(Y^r|Y>c)=\mathrm{E}\left(\frac{Yc}{\beta}\right)^r \qquad 0<\beta\leqq y \text{ all } c\geqq\beta,\ r>0$$

characterizes the Pareto distribution. Then we use these results to give generalizations of results found in some of the previously mentioned papers.

3. The results

We prove the following theorems.

THEOREM 3.1. *Let X, a non-negative random variable, with distribution function $F(x)$ and $F(a)=1$ for some $a>0$. If* (2.1) *holds for some $r>0$, then $F(x)$ is the power distribution* (1.1) *for some $\theta>0$.*

PROOF. Using the definition of the expectation, equation (2.1) is written:

$$\text{(3.1)} \qquad \int_0^c x^r d\frac{F(x)}{F(c)}=c^r\int_0^a\left(\frac{x}{a}\right)^r dF(x)\,.$$

Using integration by parts we get:

$$c^r-r\int_0^c\frac{F(x)}{F(c)}x^{r-1}dx=c^r\delta$$

or

$$F(c)c^r = \frac{r}{1-\delta}\int_0^c F(x)x^{r-1}dx \; . \tag{3.2}$$

We set $\delta = \int_0^a (x/a)^r dF(x)$. This means $0<\delta<1$ as $F(x)$ is increasing. As the integral is a continuous function of its upper limit we conclude that $F(c)$ is continuous. This in turn means that the derivative of the integral exists, hence $F(c)$ is differentiable. So differentiating (3.2) we get

$$F'(c)c^r + rc^{r-1}F(c) = \frac{r}{1-\delta}c^{r-1}F(c)$$

where $F'(c)$ is the derivative of $F(c)$. After the calculations we are led to

$$cF'(c) = \theta F(c) \tag{3.3}$$

where $\theta = r(\delta/(1-\delta)) > 0$. Integrating (3.3) we get $F(c) = Ac^\theta$ and using the condition $F(a)=1$ we have $F(c)=(c/a)^\theta$, $\theta>0$. This proves the theorem.

For the Pareto distribution we have the following

THEOREM 3.2. *Let Y a random variable with distribution function $G(y)$ and $y \geqq \beta$ for some positive β. If* (2.2) *holds for some $r>0$ then $G(y)$ is the Pareto distribution* (1.2) *with some $\gamma>0$. We assume* $\mathrm{E}(Y^r) < \infty$.

PROOF. We denote the conditional distribution of Y, given $Y>c$, by $G_c(y)$. That is

$$G_c(y) = \begin{cases} \dfrac{G(y)-G(c)}{1-G(c)} & y>c \\ 0 & y \leqq c \; . \end{cases}$$

We set also $P(y) = 1-G(y)$ and $P_c(y) = 1-G_c(y) = P(y)/P(c)$. By definition

$$\begin{aligned} \mathrm{E}(Y^r \mid Y>c) &= \int_{c+}^\infty y^r dG_c(y) \\ &= -\int_{c+}^\infty y^r dP_c(y) \\ &= -y^r \frac{P(y)}{P(c)}\Big|_c^\infty + r\int_c^\infty \frac{P(y)}{P(c)} y^{r-1}dy \\ &= c^r + r\int_c^\infty \frac{P(y)}{P(c)} y^{r-1} dy \; . \end{aligned} \tag{3.4}$$

This is because $\lim_{y\to\infty} P(y)y^r=0$ as $\mathrm{E}\,(y^r)<\infty$. Therefore equation (2.2), using (3.4), becomes

$$c^r+r\int_c^\infty \frac{P(y)}{P(c)}y^{r-1}dy=c^r\delta$$

or

$$\int_c^\infty P(y)y^{r-1}dy=\frac{c^r(\delta-1)}{r}P(c) \tag{3.5}$$

where

$$\delta=\int_\beta^\infty \left(\frac{y}{\beta}\right)^r dG(y)\geqq 1 \quad \text{as} \quad \beta>0\,.$$

From (3.5), as before, we conclude that $P(c)$ is differentiable. So we have the differential equation

$$cP'(c)=-\gamma P(c) \tag{3.6}$$

with $\gamma=r\delta/(\delta-1)>0$. From this equation we conclude that $P(c)=Ac^{-\gamma}$, with A a constant. This means that $P(y)$ must be of the form $P(y)=d(\beta/y)^r$ with $y\geqq\beta$ and d is a constant, $0\leqq d\leqq 1$. Hence

$$G(y)=\begin{cases} 1-d\left(\dfrac{\beta}{y}\right)^r & y\geqq\beta \\ 0 & y<\beta\,. \end{cases} \tag{3.7}$$

If $d\neq 1$, then $G(y)$ has a discontinuity at β. The jump is $1-d$. But d cannot be arbitrary. Before we set $\delta=\mathrm{E}\,(y/\beta)^r$ or $\mathrm{E}\,(y/\beta)^r=\beta^r\gamma/(\gamma-r)$. So setting $G(y)$ in this, we get after some calculations that $d=1$. This proves the theorem.

If instead of equation (2.2) we consider the relation

$$\mathrm{E}\,(Y^r\,|\,Y>c)=bc^r\,, \qquad b \text{ constant independent of } c,\ r>0 \tag{3.8}$$

then the solution is equation (3.7). That is we have a distribution with a positive probability mass at β and the rest is the Pareto distribution. The probability mass is evaluated from (3.7).

The characterizations now follow from the theorems and Section 1. We can prove that $\mathrm{E}\,(X^{-r}\,|\,X<c)=\mathrm{E}\,(cX/a)^{-r}$ characterizes the power distribution. We can prove this either directly or using the fact that $1/X$ has the Pareto distribution and Theorem 3.2. The same modifications can be made in Theorem 3.2 for the Pareto distribution.

In the following we apply Theorems 3.1 and 3.2 to give some generalizations of results found in some of the papers mentioned in Section 1.

4. Applications

Let $X_1 \leqq X_2 \leqq \cdots \leqq X_n$ the order statistics of a random sample, of size n, from a distribution $F(x)$. In [5] M. Fisz, for $n=2$, proves that the independence of X_1/X_2 and X_2, for $F(x)$ absolutely continuous with $F(0)=0$, characterizes the power distribution. In [7] G. Rogers uses the independence of X_i/X_{i+1} and X_{i+1}, for some i in $1 \leqq i \leqq n-1$, to prove that an absolutely continuous $F(x)$ with $F(0)=0$ is the power distribution. But in his proof, he uses the weaker assumption that X_i/X_{i+1} has a constant conditional expectation on X_{i+1}. Then, T. S. Ferguson [6], generalizes Rogers's result by considering only continuous $F(x)$. In his paper, the relation $\mathrm{E}\,(X_i \mid X_{i+1}=x)=ax+b$ is used to characterize several distributions.

As applications we have the following corollaries:

COROLLARY 4.1. *Let* $X_i < X_{i+1}$, *order statistics from a continuous distribution* $F(x)$ *with* $F(0)=0$ *and* $F(a)=1$. *If*

$$\mathrm{E}\,(X_i^r/X_{i+1}^r \mid X_{i+1}=x)=c\,, \qquad c \text{ does not depend on } x, \text{ and } r>0 \tag{4.1}$$

holds for all x *in* $(0, a)$ *and some* i *with* $1 \leqq i \leqq n-1$, *then* $F(x)$ *is the power distribution and conversely. The same also holds for*

$$\mathrm{E}\,(X_{i+1}^r/X_i^r \mid X_{i+1}=x)=c\,, \qquad c \text{ independent of } x \text{ and } r>0 \tag{4.2}$$

provided that the expectation is finite.

PROOF. We have that the distribution of X_i given $X_{i+1}=x$ is the same as the distribution of the maximum of a sample of size i from the distribution $F(y)/F(x)$, $0 \leqq y \leqq x$. This follows from the definition of the order statistics. Hence $\Pr(X_i \leqq y \mid X_{i+1}=x)=\{F(y)/F(x)\}^i$ and because of this equation (4.1) becomes

$$\int_0^x y^r d\left\{\frac{F(y)}{F(x)}\right\}^i = cx^r\,. \tag{4.3}$$

As $0 < X_i < X_{i+1}$, from (4.1) we have that $0 \leqq c \leqq 1$. But then equation (4.3) is the same as (3.1) with $(F(x))^i$ instead of $F(x)$. From this we conclude that $F(x)$ is the power distribution with $\theta=r\delta/(i(1-\delta))>0$. In a similar way or using the observation made at the end of Section 3, we can prove the same for equation (4.2). When $F(x)$ is the power distribution, a direct calculation shows that both (4.1) and (4.2) hold. Hence the corollary is proved.

Analogous corollary holds for the Pareto distribution.

COROLLARY 4.2. *Let* $X_i < X_{i+1}$ *order statistics from a continuous* *distribution* $G(x)$ *with* $G(\beta)=0$ *for some* $\beta>0$. *If*

(4.4) $$\mathrm{E}\,(X_{i+1}^r/X_i^r\,|\,X_i=x)=c\,, \qquad c \text{ independent of } x,\ r>0$$

holds for all $x>\beta$, *then* $G(x)$ *is the Pareto distribution, provided that the expectation is finite. The same holds for the relation*

(4.5) $$\mathrm{E}\,(X_i^r/X_{i+1}^r\,|\,X_i=x)=c\,, \qquad c \text{ independent of } x,\ r>0\,.$$

Conversely, if $G(x)$ *is the Pareto distribution,* (4.4) *and* (4.5) *hold.*

PROOF. The conditional distribution of X_{i+1} given $X_i=x$ is the same as the distribution of the minimum of $n-i$ observations from the distribution $G_x(y)$. Also $c\geqq 1$. So we arrive at an equation similar to (3.5) with $(P(x))^{n-i}$ instead of $P(x)$. Hence the result.

Consider now [10]. In this paper M. Srivastava proves the following: "For an absolutely continuous distribution $F(x)$ with $F(0)=0$, the independence of $(X_1+X_2+\cdots+X_n)/X_n$ and X_n characterizes the power distribution. Correspondingly the independence of X_1 and $(X_1+X_2+\cdots+X_n)/X_1$ characterizes the Pareto distribution." Here we have the following corollary.

COROLLARY 4.3. *Let* $X_1<X_2<\cdots<X_n$ *the order statistics from a continuous distribution* $F(x)$ *with* $F(0)=0$, $F(a)=1$. *Then*

(4.6) $$\mathrm{E}\left(\sum_{i=1}^n \frac{X_i^r}{X_n^r}\,\middle|\,X_n=x\right)=c\,, \qquad c \text{ independent of } x,\ r>0$$

iff $F(x)$ *is* (1.1) *with some* $\theta>0$.

PROOF. If (4.6) holds, as $X_n>X_i>0$, then $n>c>1$. The distribution $X_1,\cdots,X_{n-1}$ given $X_n=x$ is the same as the distribution of the order statistics of a random sample of size $n-1$ from $F(y)/F(x)$ $0\leqq y\leqq x$, if $F(x)$ is continuous. But the sum of the order statistics equals the sum of the sample. Hence we have for equation (4.6)

$$\mathrm{E}\left(\sum_{i=1}^n X_i^r\,|\,X_n=x\right)=(c-1)x^r$$

or

$$\mathrm{E}\,(Y^r)=\frac{c-1}{n-1}x^r$$

where Y has $F(y)/F(x)$ as distribution. Therefore (4.6) is equivalent to

(4.7) $$\int_0^x y^r d\frac{F(y)}{F(x)}=\frac{c-1}{n-1}x^r\,.$$

As $0<(c-1)/(n-1)<1$ equation (4.7) is the same like (3.1). Hence the result.

For the Pareto distribution we state an analogous corollary

COROLLARY 4.4. *The assumptions as in Corollary* 4.3, *with* $F(\beta)=0$ *for some* $\beta>0$. *Then*

$$\mathrm{E}\left(\sum_{i=1}^{n}\frac{X_i^r}{X_1^r}\bigg|X_1=x\right)=c\,,\qquad c \text{ independent of } x,\ r>0 \tag{4.8}$$

iff $F(x)$ *is* (1.2) *with some* $\gamma>0$. *We assume that the expectation is finite.*

PROOF. On the same lines as in Corollary 4.3.

ATHENS UNIVERSITY

REFERENCES

[1] Basu, A. P. (1965). On characterizing the exponential distribution by order statistics, *Ann. Inst. Statist. Math.*, **17**, 93-96.

[2] Desu, M. M. (1971). A characterization of the exponential distribution by order statistics, *Ann. Math. Statist.*, **42**, 837-838.

[3] Govindarajulu, Z. (1966). Characterizations of the exponential and power distributions, *Skand. Aktuarietidskr.*, **49**, 132-136.

[4] Rossberg, H. J. (1966). Charakterisierungs probleme, die sich aus der von A. RENYI in die Theorie der Ranggrössen eingeführten Methode ergeben, *Monat. Deuts. Akad. Wiss. Berlin*, 8, 561-572.

[5] Fisz, M. (1958). Characterizations of some probability distributions, *Skand. Aktuarietidskr.*, **41**, 65-67.

[6] Ferguson, T. S. (1967). On characterizing distributions by properties of order statistics, *Sankhyā* A, **29**, 265-278.

[7] Rogers, G. S. (1963). An alternative proof of the characterization of the density $A\times B$, *Amer. Math. Monthly*, **70**, 857-858.

[8] Malik, H. J. (1970). A characterization of the Pareto distribution, *Skand. Aktuartidskr.*, **53**, 115-117.

[9] Krishnaji, N. (1970). Characterization of the Pareto distribution through a model of underreported incomes, *Econometrica*, **38**, 251-255.

[10] Srivastava, M. S. (1965). A characterization of Pareto's distribution and $(k+1)x^k/\vartheta^{(k+1)}$, (Abstract), *Ann. Math. Statist.*, **36**, 361-362.