Physica A xx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Q1 A new Pareto-type distribution with applications in reliability and income data

Q2 Marcelo Bourguignon [a,*], Helton Saulo [b], Rodrigo Nobre Fernandez [c]

[a] *Departamento de Estatística, Universidade Federal do Rio Grande do Norte, Natal, RN, Brazil*
[b] *Instituto de Matemática e Estatística, Universidade Federal de Goiás, Goiânia, Brazil*
[c] *Department of Economics, Universidade Federal de Pelotas, Pelotas, Brazil*

HIGHLIGHTS

- The model is characterized by transformation of the half logistic distribution.
- The proposed model has only two parameters.
- The model presents more simplicity in mathematical terms.

ARTICLE INFO

Article history:
Received 20 July 2015
Received in revised form 16 February 2016
Available online xxxx

Keywords:
Pareto distribution
Income distribution
Maximum likelihood estimation
Gini coefficient
Reliability

ABSTRACT

A new Pareto-type distribution is introduced and studied. This new model is a generalization of the well-known Pareto distribution. We derive some of its probabilistic and inferential properties. We deduce the mathematical form of the Lorenz curve and the Gini index associated with the new model. The maximum likelihood estimators are derived and their performance are evaluated through a Monte Carlo simulation study. Finally, we illustrate the flexibility of the new distribution by means of three applications to real data sets.

1. Introduction

The Pareto distribution (Pareto [1]) was originally proposed to model the unequal distribution of wealth since he observed the way that a larger portion of the wealth of any society is owned by a smaller percentage of the people. The Pareto distribution plays an important role in analyzing a wide range of real-world situations, not only in the field of economics. For example, in actuarial sciences, reliability, finance and climatology, where it usually describes the occurrence of extreme weather. Several alternative forms of the Pareto distribution can be found in the literature. Dragulescu and Yakovenko [2], Silva [3], Yakovenko and Rosser [4] advanced in an exponential distribution of individual income similar to the Boltzmann–Gibbs. Clementi et al. [5–7] proposed a generalized κ distribution and Willis and Mimkes [8] used lognormal and Boltzmann distributions to argue for a separate treatment of employed and self-employed income. Moura Jr. and Ribeiro [9] and Figueira et al. [10] showed that the Gompertz curve combined with Pareto's law is a good descriptive model for income distribution.

In this context, this paper aims to introduce a new Pareto-type distribution that offers a better fit in certain practical situations (see Section 7). The new distribution is characterized by transformation of the half logistic distribution. We

* Corresponding author.
E-mail address: m.p.bourguignon@gmail.com (M. Bourguignon).

http://dx.doi.org/10.1016/j.physa.2016.03.043
0378-4371/

provide a comprehensive account of the mathematical properties of the proposed distribution. It should be highlighted that, compared with the Pareto distribution, our model has the following advantages: (1) it can have an upside-down bathtub or a decreasing hazard rate function depending on the values of its parameters; (2) mathematical simplicity. For instance, the probability and distribution functions of the new distribution have a simple form in contrast with some generalizations of the Pareto distribution which have associated probability function involving special functions (log, beta and gamma functions); and (3) the proposed distribution has only two parameters, in contrast with some generalizations of the Pareto distribution which have three or four parameters.

This study is organized in eight sections starting with this introduction. In Sections 2 and 3, we present the new model and derive some of its properties. Explicit expressions for the Lorenz curve and Gini index are derived in Section 4. In Section 5, we discuss maximum likelihood (ML) estimation of the model parameters. We present a Monte Carlo (MC) simulation experiment to evaluate the ML estimates of the model parameters in Section 6. Three applications in Section 7 illustrate the usefulness of the new distribution for data modeling. Finally, we have the final considerations.

2. The new model

The cumulative distribution function (CDF) of the new Pareto-type distribution (NP) is given by

$$F(x;\alpha,\beta)=\frac{1-(\beta/x)^{\alpha}}{1+(\beta/x)^{\alpha}}=\frac{x^{\alpha}-\beta^{\alpha}}{x^{\alpha}+\beta^{\alpha}}=1-\frac{2\,(\beta/x)^{\alpha}}{1+(\beta/x)^{\alpha}}=1-\frac{2\,\beta^{\alpha}}{x^{\alpha}+\beta^{\alpha}},\quad x\geq\beta, \tag{1}$$

where $\alpha>0$ is a shape parameter and $\beta>0$ is the scale parameter. In this case, the notation $X\sim \mathrm{NP}(\alpha,\beta)$ is used.

For high incomes this formula closely approximates the form

$$F(x;\alpha,\beta)=1-A\,x^{-\alpha},$$

which is the form predicted by Pareto's law. The model specified in (1) possesses this important property of the weak Pareto law (Mandelbrot [11]), that is, (1) converges in distribution to the Pareto model for x sufficiently large.

Lemma 1 (*A Characterization*)**.** *If a random variable (RV) Y follows the half logistic distribution with parameters α, then the RV $X=\beta\,\mathrm{e}^{Y}$ has the CDF in Eq.* (1) *with parameters α and β.*

The representation above highlights the useful observation that the logarithm of such a variable has a shifted half logistic distribution. It will permit the recognition of many distributional properties of the new model as reflections of parallel properties of half logistic variables.

Remark 1. If Z has the exponential distribution then $Y=\beta\mathrm{e}^{Z}$ has the Pareto distribution.

Remark 2. If Z has the logistic distribution then $Y=\beta\mathrm{e}^{Z}$ has the log–logistic distribution.

The probability density function (PDF) corresponding to (1) is

$$f(x;\alpha,\beta)=\frac{2\,\alpha\,(\beta/x)^{\alpha}}{x[1+(\beta/x)^{\alpha}]^{2}}=\frac{2\,\alpha\,(\beta/x)^{\alpha+1}}{\beta[1+(\beta/x)^{\alpha}]^{2}}=\frac{2\,\alpha\,\beta^{\alpha}x^{\alpha-1}}{(x^{\alpha}+\beta^{\alpha})^{2}},\quad x\geq\beta. \tag{2}$$

It can be shown that

$$\lim_{x\to\beta}f(x;\alpha,\beta)=\frac{\alpha}{2\,\beta}\quad\text{and}\quad\lim_{x\to\infty}f(x;\alpha,\beta)=0.$$

Proposition 1. *For $x\geq\beta>0$, the NP density function is decreasing.*

Proof. The first derivative of $f(x;\alpha,\beta)$ becomes

$$f'(x;\alpha,\beta)=\frac{\mathrm{d}\,[f(x;\alpha,\beta)]}{\mathrm{d}x}=\frac{2\,\alpha\,\beta^{a}\,x^{\alpha-2}}{(x^{\alpha}+\beta^{\alpha})^{3}}\Delta(x),$$

where $\Delta(x)=(\alpha-1)\beta^{\alpha}-(\alpha+1)x^{\alpha}$. Clearly, $\Delta(x)$ is strictly decreasing in x with $\Delta(\beta)=-2\,\beta^{\alpha}$ and $\Delta(\infty)=-\infty$.

Remark 3. The density function of X can be expressed as

$$f(x;\alpha,\beta)=\frac{2}{[1+(\beta/x)^{\alpha}]^{2}}\cdot g(x;\alpha,\beta),$$

where $g(x;\alpha,\beta)=\alpha\,\beta^{\alpha}x^{-(\alpha+1)}$ is the PDF of the Pareto distribution. The multiplying quantity $2/[1+(\beta/x)^{\alpha}]^{2}$ works as a corrected factor for the PDF of the Pareto model.

Please cite this article in press as: M. Bourguignon, et al., A new Pareto-type distribution with applications in reliability and income data, Physica A (2016), http://dx.doi.org/10.1016/j.physa.2016.03.043