Available online at www.sciencedirect.com
ScienceDirect
Journal of Informetrics 2 (2008) 156–163
Journal of INFORMETRICS
An International Journal
www.elsevier.com/locate/joi

A general definition of the Leimkuhler curve

José María Sarabia *
Department of Economics, University of Cantabria, Avda. de los Castros s/n, 39005 Santander, Spain
Received 15 October 2007; received in revised form 16 January 2008; accepted 17 January 2008

Abstract

In this paper, we provide a general definition of the Leimkuhler curve in terms of the theoretical cumulative distribution function. The definition applies to discrete, continuous and mixed random variables. Several examples are given to illustrate the use of the formula.

Keywords: Productivity; Cumulative distribution function; Incomplete beta function; Lorenz curve

1. Introduction

The Lorenz and Leimkuhler curves are two important instruments in informetrics (Burrell, 1991, 2006, 2007; Egghe, 2005; Rousseau, 2007). These curves plot the cumulative proportion of total productivity against the cumulative proportion of sources. The difference between the two constructions is that for the Lorenz curve in economics, one arranges the sources in increasing order of productivity while for the Leimkuhler curves in informetrics they are arranged in decreasing order.

In this paper, we provide a general definition of the Leimkuhler curve in terms of the theoretical cumulative distribution function (cdf). The definition applies to discrete, continuous and mixed random variables.

The contents of this paper are as follows. In Section 2, we propose a general definition of the Leimkuhler curve. Section 3 shows the equivalence with the classical definition. Finally, several illustrative examples are presented in Section 4, including exponential, Singh–Maddala, power, arcsine, lognormal, mixed Pareto and geometric distributions.

2. The general definition

Let X be a random variable, which denotes the productivity of a randomly chosen source. We assume that X has cumulative distribution function $F_X(x)$, which represents the proportion of the population with productivity less than or equal to x.

Now, let $\mathcal{L}$ be the class of all non-negative random variables with positive finite expectations. For a random variable X in $\mathcal{L}$ with cumulative distribution function F_X we define its inverse distribution function by

$$F_X^{-1}(y) = \inf\{x \; : \; F_X(x) \geq y\}. \tag{1}$$

* Corresponding author. Tel.: +34 942 201635; fax: +34 942 201603.
E-mail address: sarabiaj@unican.es.

1751-1577/$ – see front matter
doi:10.1016/j.joi.2008.01.002

The mathematical expectation of X is

$$\mu_X = \int_0^\infty x\,\mathrm{d}F_X(x) = \int_0^1 F_X^{-1}(y)\mathrm{d}y,$$

after making the change of variable $y = F_X(x)$, and where the last expression is given in terms of the inverse of the cumulative distribution function. Next definition makes use of the inverse of the cdf, as in the Gastwirth's (1971) definition of the Lorenz curve. However, the Gastwirth's formulation has not previously been used in an informetrics context.

Definition 1. Let $X \in \mathcal{L}$ with cumulative distribution function F_X and inverse distribution function F_X^{-1}. The Leimkuhler curve K_X corresponding to X is defined by

$$K_X(p) = \frac{1}{\mu_X}\int_{1-p}^1 F_X^{-1}(y)\mathrm{d}y, \quad 0 \le p \le 1. \tag{2}$$

This definition contains the cases of discrete, continuous and mixed random variables. In formula (2)F_X^{-1} is piecewise continuous and the integrals can be assumed to be ordinary Riemann integrals.

From definition (2) we can show that a Leimkuhler curve will be a continuous, non-decreasing concave function that is differentiable almost everywhere in [0, 1] with $K_X(0) = 0$ and $K_X(1) = 1$.

The Leimkhuler curve determines the distribution of X up to a scale factor transformation. This is true since $F_X^{-1}(x) = \mu_X K_X'(1-x)$, almost everywhere and F_X^{-1} will determine F_X.

There is a simple relation between the Leimkuhler curves of the random variables X and $Y = X + \lambda$, where $\lambda \ge 0$.

Proposition 1. *Let X be random variable in $\mathcal{L}$ with mathematical expectation μ_X and Leimkuhler curve $K_X(p)$. Then, the Leimkuhler curve of the random variable $Y = X + \lambda$ is,*

$$K_Y(p) = \frac{\lambda p + \mu_X K_X(p)}{\lambda + \mu_X}, \tag{3}$$

which is a convex linear combination of the Leimkuhler curves p and $K_X(p)$.

Proof. Let $F_X(\cdot)$ be the cdf of X. Then, $F_Y(y) = F_X(y-\lambda)$, $F_Y^{-1}(y) = \lambda + F_X^{-1}(y)$ and $E(Y) = \lambda + \mu_X$. Now, using (2)

$$K_Y(p) = \frac{1}{\mu_Y}\int_{1-p}^1 F_Y^{-1}(y)\mathrm{d}y = \frac{\lambda p + \int_{1-p}^1 F_X^{-1}(y)\mathrm{d}y}{\lambda + \mu_X},$$

which is (3). □

Finally, we include the relation between the Lorenz and Leimkuhler curves. If $L_X(p)$ is a Lorenz curve, it is direct to show that,

$$K_X(p) = 1 - L_X(1-p), \quad 0 \le p \le 1.$$

3. Equivalence with the classical definition

The usual definition of Leimkuhler curve (see for example Burrell, 1991, 1992) is based on two equations. First one solve for x,

$$p = 1 - F_X(x) = \int_x^\infty f_X(y)\mathrm{d}y, \tag{4}$$

and then considers

$$K_X(p) = \psi_X(x) = \frac{1}{\mu_X}\int_x^\infty y f_X(y)\mathrm{d}y. \tag{5}$$

If we make the change of variable

$$u = F_X(y),$$

a solution of this equation is (1) and (5) becomes in

$$K_X(p) = \frac{1}{\mu_X} \int_{F_X(x)}^{1} F_X^{-1}(u)\mathrm{d}u,$$

which is formula (2). Note that from (4)$F_X(x) = 1 - p$.

An alternative, equivalent, Gastwirth-style general definition can also be given in terms of the inverse of the tail distribution function $Pr(X \geq x)$.

4. Examples

In this section, several examples are given in order to illustrate the use of formula (2).

4.1. Two-parameter exponential distribution

We begin with the one-parameter exponential distribution with cdf,

$$F_X(x) = 1 - \exp\left(\frac{-x}{\sigma}\right), \quad x \geq 0, \tag{6}$$

and $F_X(x) = 0$ if $x < 0$, where $\sigma > 0$. For distribution (6) we have $\mu_X = \sigma$ and

$$F_X^{-1}(y) = -\sigma \log(1 - y), \quad 0 < y < 1.$$

Then, using (2) we obtain,

$$K_X(p) = \frac{1}{\mu_X} \int_{1-p}^{1} \left[-\sigma \log(1 - y)\right] \mathrm{d}y = p - p \log p. \tag{7}$$

Now, let Y be a two-parameter exponential distribution with cdf,

$$F_Y(y) = 1 - \exp\left[\frac{-(y - \lambda)}{\sigma}\right], \quad y \geq \lambda, \tag{8}$$

and $F_Y(y) = 0$ if $y < \lambda$, where $\sigma > 0$ and $\lambda \geq 0$. An interesting informetrics application in the case where $\lambda = 1$ is given in Lafouge (2007). It is satisfies that $Y = X + \lambda$. Then, using (3) and (7) we have,

$$K_Y(p) = \frac{\lambda p + \sigma K_X(p)}{\lambda + \sigma} = \frac{p(\lambda + \sigma - \sigma \log p)}{\lambda + \sigma},$$

which is the Leimkuhler curve corresponding to (8).

4.2. Singh–maddala distribution

The Singh–Maddala distribution is one of the most popular distributions used in practice to fit income and wealth data. This distribution was obtained by Singh and Maddala (1976) by considering the hazard rate of income. Let X be a random variable with Singh–Maddala distribution with cdf,

$$F_X(x) = 1 - \frac{1}{[1 + (x/\sigma)^a]^q}, \quad x > 0 \tag{9}$$

where $a, q, \sigma > 0$. Definition (9) corresponds to the Pareto IV distribution, in the Arnold (1983) Pareto hierarchy. If we set $q = 1$ in (9), we obtain the Fisk or log–logistic distribution, which has been used in analysis of citation age data, or aging of information sources (see Burrell, 2002; Basulto and Ortega, 2005).

We have next result: the Leimkuhler curve of the Singh–Maddala distribution is given by

$$K_X(p) = 1 - I_{1-p^{1/q}}\left(\frac{1}{a} + 1, q - \frac{1}{a}\right), \tag{10}$$

if $q > 1/a$, where $I_x(,)$ represents the incomplete beta function ratio defined as

$$I_x(a, b) = \int_0^x \frac{t^{a-1}(1-t)^{b-1}}{B(a, b)}\mathrm{d}t, \quad 0 < x < 1, \tag{11}$$

where $B(a, b) = \int_0^1 t^{a-1}(1-t)^{b-1}\,\mathrm{d}t$ denotes the classical beta function, where $a, b > 0$. Note that in probability and statistics (11) corresponds to the cumulative distribution function of the beta distribution of the first kind.

If $r = 1/a + 1$ and $s = q - 1/a$ is a positive integer (10) becomes in,

$$K_X(p) = 1 - \sum_{i=0}^{s-1}\left[\frac{\Gamma(q+1)}{\Gamma(r+i-1)\Gamma(s-i)}\right](1-p^{1/q})^{r+i} p^{(s-i+1)/q}, \tag{12}$$

and if both r and s are positive integers we have,

$$K_X(p) = 1 - \sum_{j=r}^{q}\binom{q}{j}(1-p^{1/q})^j p^{(q-j)/q}. \tag{13}$$

The proof of (10) is the following. The inverse of the cdf of the Singh–Maddala distribution is

$$F_X^{-1}(y) = \sigma[(1-y)^{-1/q} - 1]^{1/a}, \quad 0 \leq y \leq 1.$$

We assume $q > 1/a$ in order to have a finite mean. Using (2) and making the change of variable $t = 1 - (1-y)^{1/q}$ we obtain,

$$\begin{aligned}
K_X(p) &= \frac{1}{\mu_X}\int_{1-p}^{1} \sigma[(1-y)^{-1/q} - 1]^{1/a}\,\mathrm{d}y \\
&= \frac{\sigma q}{\mu_X}\int_{1-p^{1/q}}^{1} t^{1/a}(1-t)^{q-1/a-1}\,\mathrm{d}t \\
&= \frac{\sigma q}{\mu_X}\left\{B(1/a+1, q-1/a)\left[1 - I_{1-p^{1/q}}(1/a+1, q-1/a)\right]\right\} \\
&= 1 - I_{1-p^{1/q}}(1/a+1, q-1/a).
\end{aligned}$$

Finally, taking in consideration next two formulas: if b is a positive integer and $a \geq 0$:

$$I_x(a, b) = \sum_{i=0}^{b-1}\left[\frac{\Gamma(a+b)}{\Gamma(a+i+1)\Gamma(b-i)}\right] x^{a+i}(1-x)^{b-i-1},$$

and if a and b are both positive integers we have,

$$I_x(a, b) = \sum_{j=a}^{a+b-1}\binom{a+b-1}{j} x^j(1-x)^{a+b-1-j},$$

we obtain formulas (12) and (13), respectively. The incomplete beta function ratio can be computed with standard software like Excel or Mathematica.

A simpler formula can be obtained when $q = 1 + 1/a$. If $a > 0$, $b = 1$ and $0 \leq x \leq 1$,

$$I_x(a, 1) = \int_0^x \frac{t^{a-1}}{B(a, 1)}\mathrm{d}t = x^a. \tag{14}$$

Using now formula (14), general expression (10) becomes in:

$$K_X(p) = 1 - I_{1-p^{a/(a+1)}}\left(\frac{1}{a}+1, 1\right)$$
$$= 1 - [1 - p^{a/(a+1)}]^{1/a+1}.$$

The corresponding Lorenz curve is a particular case of the general family proposed by Sarabia, Castillo and Slottje (1999).

4.3. Power distribution

Let X be a random variable with cdf

$$F_X(x) = \left(\frac{x}{\sigma}\right)^{\alpha}, \quad 0 \le x \le \sigma, \tag{15}$$

with $\alpha, \sigma > 0$. This distribution is called a power distribution (see (Balakrishnan & Nevzorov, 2003). In the special case $\alpha = 1$, we have a uniform distribution in the interval $[0, \sigma]$. Distribution (15) can be used as a possible model in analysis of citation age data. The mathematical expectation of (15) is $\mu_X = \alpha\sigma/(\alpha+1)$ and the the inverse of the cdf,

$$F_X^{-1}(y) = \sigma y^{1/\alpha}, \quad 0 \le y \le 1.$$

The Leimkuhler of X is,

$$K_X(p) = \frac{1}{\mu_X}\int_{1-p}^{1}\left(\sigma y^{1/\alpha}\right)\mathrm{d}y = 1 - (1-p)^{1/\alpha+1}. \tag{16}$$

If we include a location parameter in (15) we obtain,

$$F_Y(y) = \left(\frac{y-\lambda}{\sigma}\right)^{\alpha}, \quad \lambda \le y \le \lambda + \sigma,$$

with $\alpha, \sigma > 0$, $\lambda \ge 0$ and $Y = X + \lambda$. If we use (3) and (16) we have

$$K_Y(p) = \frac{\lambda p + \mu_X K_X(p)}{\lambda + \mu_X} = \frac{\lambda(\alpha+1)p + \alpha\sigma[1-(1-p)^{1/\alpha+1}]}{\lambda(\alpha+1)+\alpha\sigma}.$$

For a similar result in the context of Lorenz curves, see Sarabia (1997).

4.4. Arcsine distribution

We say that a random variable X has an arcsine distribution if its cdf is given by

$$F_X(x) = \frac{2}{\pi}\arcsin\sqrt{\frac{x}{\sigma}}, \quad 0 \le x \le \sigma, \tag{17}$$

with $\sigma > 0$. The arcsine distribution is a linear transformation $X = \sigma Z$, where Z is distributed according to a classical beta distribution with parameters $(1/2, 1/2)$. The importance of this distribution is due to its applications in the theory of random walks. Let us consider partial sums $S_0 = 0$, $S_k = U_1 + \cdots + U_k$, $k = 1, 2, \ldots, n$, where independent random variables $U_1, U_2, \ldots$ take on values -1 and 1 with equal probabilities. Let us define by V_n the number of positive partial sums among $S_0, S_1, \ldots, S_n$. In consequence, $W_n = V_n/n$ is the fraction of positive sums and

$$\lim_{n\to\infty} Pr(W_n \le x) = \frac{2}{\pi}\arcsin\sqrt{x}, \quad 0 \le x \le 1,$$

which corresponds to (17) with $\sigma = 1$. Now, let us consider the index of the maximal sum,

$$I_n = \min\{j : S_j = \max\{S_0, S_1, \ldots, S_n\}\},$$

and again

$$\lim_{n\to\infty} Pr(I_n/n \le x) = \frac{2}{\pi} \arcsin \sqrt{x}, \quad 0 \le x \le 1.$$

A more detail description of this distribution can be found in Arnold and Groeneveld (1980) and Johnson, Kotz and Balakrishnan (1995).

The mathematical expectation of (17) is $\mu_X = \sigma/2$. Since,

$$F_X^{-1}(y) = \sigma \sin^2(\pi y/2), \quad 0 \le y \le 1,$$

we have that,

$$K_X(p) = \frac{1}{\mu_X} \int_{1-p}^{1} \left[\sigma \sin^2(\pi y/2)\right] \mathrm{d}y = \frac{1}{\pi} \left[\pi p + \sin(\pi p)\right].$$

If we work with the translated version $Y = X + \lambda$, $\lambda \ge 0$ with cdf,

$$F_Y(y) = \frac{2}{\pi} \arcsin \sqrt{\frac{y-\lambda}{\sigma}}, \quad \lambda \le y \le \lambda + \sigma,$$

the new Leimkuhler curve is

$$K_Y(p) = \frac{\lambda p + (\sigma/2)K_X(p)}{\lambda + \sigma/2} = \frac{\lambda p + (\sigma/2\pi)[\pi p + \sin(\pi p)]}{\lambda + \sigma/2}.$$

4.5. Lognormal distribution

Let X be a lognormal distribution with cdf given by

$$F_X(x) = \Phi\left(\frac{\log x - \mu}{\sigma}\right), \quad x > 0, \tag{18}$$

where $\Phi(\cdot)$ denotes the cdf of the standard normal distribution, with $-\infty < \mu < \infty$ and $\sigma > 0$. A lognormal distribution will be denoted by $X \sim \mathcal{LN}(\mu, \sigma)$.

Proposition 2. *Let $X \sim \mathcal{LN}(\mu, \sigma)$ be a classical lognormal distribution with cdf (18). Then, the Leimkuhler curve is given by*

$$K_X(p) = \Phi(\sigma - \Phi^{-1}(1-p)), \quad 0 \le p \le 1. \tag{19}$$

Proof. Because μ in (19) is a scale parameter, we can assume $\mu = 1$. Then

$$F_X^{-1}(x) = \exp[1 + \sigma\Phi^{-1}(x)], \quad 0 \le x \le 1,$$

and $\mu_X = \exp(1 + \sigma^2/2)$. Now, using (2),

$$\begin{aligned}\int_{1-p}^{1} F_X^{-1}(y)\mathrm{d}y &= \int_{F_X^{-1}(1-p)}^{\infty} t f_X(t)\mathrm{d}t \\ &= \int_{F_X^{-1}(1-p)}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2}(\log t - 1)^2\right\} \mathrm{d}t \\ &= \int_{\log F_X^{-1}(1-p)}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2}(z-1)^2 + 1\right\} \mathrm{d}z,\end{aligned}$$

where we have made the successive changes of variable $F_X^{-1}(y) = t$ and $\log t = z$. Last integral can be written as,

$$\exp(1 + \sigma^2/2) \int_{\log F_X^{-1}(1-p)}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2}(z - (1+\sigma^2))^2\right\} \mathrm{d}z. \tag{20}$$

The kernel of the integral in (20) corresponds to the probability density function of a normal distribution with mathematical expectation $1+\sigma^2$ and variance σ^2. Then, if we define $\tilde{Z} \sim \mathcal{N}(1+\sigma^2, \sigma^2)$, expression (20) can be written as,

$$\exp(1+\sigma^2/2)Pr(\tilde{Z} > 1+\sigma\Phi^{-1}(1-p)),$$

because $\log F_X^{-1}(1-p) = 1+\sigma\Phi^{-1}(1-p)$. Finally,

$$\begin{aligned} Pr(\tilde{Z} > 1+\sigma\Phi^{-1}(1-p)) &= 1-\Phi\left(\frac{1+\sigma\Phi^{-1}(1-p)-(1+\sigma^2)}{\sigma}\right) \\ &= \Phi(\sigma-\Phi^{-1}(1-p)), \end{aligned}$$

and we get (19). □

Formula (19) is very simple and convenient for computing, and has not been described previously in the informetrics literature.

The formula for the Lorenz curve of the lognormal distribution can be found in Arnold (1987, p. 42) or Sarabia (2008). On the other hand $K_X(p)$ is an increasing function in σ, and then if $X_i \sim \mathcal{LN}(\mu_i, \sigma_i)$, $i = 1, 2$ with $\sigma_1 \geq \sigma_2$, then $K_{X_1}(p) \geq K_{X_2}(p)$ for all $p \in (0, 1)$.

4.6. A mixed Pareto distribution

In this example, we study a mixture distribution with a discrete part and a continuous part, considered by Gastwirth (1971). Let X be a random variable in which a proportion p_0 of the population has a productivity $\sigma > 0$ and the remaining proportion $1-p_0$ of the population has a productivity distributed according to a classical Pareto distribution with parameters σ and $\alpha > 1$. The cdf of X is $F_X(x) = 0$ if $x < \sigma$, $F_X(x) = p_0$ if $x = \sigma$ and

$$F_X(x) = 1-(1-p_0)\left(\frac{x}{\sigma}\right)^{-\alpha}, \quad \text{if } p_0 < x < \infty.$$

The inverse of F_X is

$$F_X^{-1}(y) = \begin{cases} \sigma & \text{if } 0 < y < p_0 \\ \sigma\left(\dfrac{1-p_0}{1-y}\right)^{1/\alpha} & \text{if } p_0 < y < 1. \end{cases}$$

Using definition (2) we obtain the Leimkuhler curve,

$$K_X(p) = \begin{cases} \dfrac{\alpha(1-p_0)^{1/\alpha}}{\alpha-p_0}p^{1-1/\alpha} & \text{if } 0 < p < 1-p_0 \\ 1-\dfrac{(\alpha-1)(1-p)}{\alpha-p_0} & \text{if } 1-p_0 < p < 1. \end{cases}$$

If $p_0 = 0$ then $K_X(p) = p^{1-1/\alpha}$, which is the Leimkuhler curve corresponding to a classical Pareto distribution (Burrell, 1992).

4.7. A discrete example

In order to complete this section we include the Leimkuhler curve corresponding to a discrete random variable. We consider a geometric distribution. This case has been considered by Gastwirth (1971) for Lorenz curves and by Burrell (1992) in an informetric context with other different version of the geometric distribution. Let X be a geometric distribution with probability mass function $Pr(X = k) = pq^{k-1}$, $k = 1, 2, \ldots$ with $0 < p < 1$ and $q = 1-p$. The cumulative distribution function is $F_X(k) = Pr(X \leq k) = 1-q^k$. Then

$$F_X^{-1}(u) = k, \quad 1-q^{k-1} \leq u \leq 1-q^k.$$

Using formula (2), the Leimkuhler is

$$K_X(u) = kq^{k-1} - (k-1)q^k + kp(u - q^{k-1}), \quad \text{if}\, q^k \le u \le q^{k-1}.$$

Acknowledgments

The author thanks to Ministerio de Educación y Ciencia (projects SEJ2004-02810 and SEJ2007-65818) for partial support of this work. I am grateful for the constructive suggestions provided by two referees, which improved the paper.

References

Arnold, B. C. (1983). *Pareto distributions*. Fairland, Maryland: International Co-operative Publishing House.
Arnold, B. C. (1987). *Majorization and the Lorenz order: A brief introduction. Lecture notes in statistics* (Vol. 43). New York: Springer-Verlag.
Arnold, B. C., & Groeneveld, R. A. (1980). Some properties of the arcsine distribution. *Journal of the American Statistical Association*, *75*, 173–175.
Balakrishnan, N., & Nevzorov, V. B. (2003). *A primer on statistical distributions*. New York: John Wiley.
Basulto, J., & Ortega, F. J. (2005). Modelling citation age data with right censoring. *Scientometrics*, *62*, 329–342.
Burrell, Q. L. (1991). The Bradford distribution and the Gini index. *Scientometrics*, *21*, 181–194.
Burrell, Q. L. (1992). The Gini index and the Leimkuhler curve for bibliometric processes. *Information Processing and Management*, *28*, 19–33.
Burrell, Q. L. (2002). Modelling citation age data: simple graphical methods from reliability theory. *Scientometrics*, *55*, 273–285.
Burrell, Q. L. (2006). On Egghe's version of continuous concentration theory. *Journal of the American Society for Information Science and Technology*, *57*, 1406–1411.
Burrell, Q. L. (2007). Egghe's construction of Lorenz curves resolved. *Journal of the American Society for Information Science and Technology*, *58*, 2157–2159.
Egghe, L. (2005). Zipfian and Lotkaian continuous concentration theory. *Journal of the American Society for Information Science and Technology*, *56*, 935–945.
Gastwirth, J. L. (1971). A general definition of the Lorenz curve. *Econometrica*, *39*, 1037–1039.
Johnson, N. L., Kotz, S., & Balakrishnan, N. (1995). *Continuous univariate distributions* (Vol. 2, 2nd ed.). New York: John Wiley.
Lafouge, T. (2007). The source-item coverage of the exponential function. *Journal of Informetrics*, *1*, 59–67.
Rousseau, R. (2007). On Egghe's construction of Lorenz curves. *Journal of the American Society for Information Science and Technology*, *58*, 1551–1552.
Sarabia, J. M. (1997). A hierarchy of Lorenz curves based on the generalized Tukey's Lambda distribution. *Econometric Reviews*, *16*, 305–320.
Sarabia, J. M. (2008). Parametric Lorenz curves: models and applications. In D. Chotikapanich (Ed.), *Modeling income distributions, Lorenz, curves, series: Economic studies in inequality, social exclusion, well-being* (Vol. 4). Springer.
Sarabia, J. M., Castillo, E., & Slottje, D. J. (1999). An ordered family of Lorenz curves. *Journal of Econometrics*, *91*, 43–60.
Singh, S. K., & Maddala, G. S. (1976). A Function for the size distribution of incomes. *Econometrica*, *44*, 963–970.