Inequality Measures for a Mixture of Pareto Lognormal Distributions

William Griffiths*
University of Melbourne

Duangkamon Chotikapanich
Monash University

Gholamreza Hajargasht
Swinburne Business School

6 August, 2020

Abstract

Formulas are derived for the Gini, Theil and Pietra coefficients for a population-weighted mixture of Pareto-lognormal distributions, and applied to South America for three years. The formulas are useful for measuring regional or global inequality in large-scale projects that utilise Pareto-lognormal distributions.

Keywords: Gini coefficient; Theil index; Pietra index.

*Corresponding author
William Griffiths
Department of Economics
University of Melbourne
Vic 3010
Australia
wegrif@unimelb.edu.au

1. Introduction

The Pareto-lognormal distribution, a modification of the lognormal distribution which exhibits Pareto behaviour in the right tail, was suggested by Colombi (1990) for modelling income distributions. It was extended to a double Pareto-lognormal distribution by Reed (2003) and Reed and Jorgenson (2004) to capture Pareto behaviour in both tails. Several applications of the double Pareto-lognormal distribution followed. Reed (2002) and Giesen et al (2010) used it for modelling the sizes of human settlements and cities. Concerned that the assumption of a constant elasticity inverse demand function relies on the possibly unrealistic assumption of a Pareto income distribution, Fabinger and Weyl (2018) explored the double Pareto lognormal as an alternative. Hajargasht and Griffiths (2013) described how to estimate the Pareto lognormal and double Pareto-lognormal distributions using grouped data and provided expressions for calculating inequality and poverty measures from their parameters.

Our focus is on the Pareto-lognormal distribution and its use in large-scale projects designed to investigate regional or global income distributions, and the corresponding regional or global inequality. In such studies, it is often convenient to estimate parametric income distributions for the component countries and to use a population-weighted mixture of those distributions to study the combined distribution and its characteristics. Inequality measures from the mixture distribution, reflecting both within-country and between-country inequality, can be computed from the parameters of the component distributions. An example is the study by Chotikapanich et al., (2012). They estimated component-country beta-2 distributions and used their parameters to estimate global and regional inequalities. With its general lognormal shape and a tail that exhibits Pareto behaviour, the Pareto-lognormal distribution is another good candidate for estimating the component-country distributions. The purpose of this note to provide expressions that can be used to compute regional or global inequality for a mixture distribution whose components are Pareto-lognormal distributions. Formulas are provided for the Gini, Theil and Pietra coefficients as functions of the parameters of the Pareto-lognormal components. The more general double Pareto-lognormal distribution may have been a more natural choice. However, given that Pareto behaviour in the right tail, not the left tail, is the prime concern of income distribution modellers, we opted to avoid the extra algebraic complexity of the double Pareto-lognormal distribution. Hajargasht and Griffiths (2013) estimated Pareto lognormal, double Pareto lognormal, and

GB2 distributions for 10 countries and found no clear ranking in terms of goodness-of-fit. The fit of the Pareto-lognormal was comparable to the other two distributions in four of the ten cases. The GB2 distribution is generally a popular choice,[1] but, in any event, when undertaking large scale projects involving parametric distributions, it is prudent to investigate the sensitivity of inequality estimates to alternative distributional assumptions.

In Section 2 we specify the density and distribution functions for the mixture of Pareto-lognormal distributions. Expressions for the Gini, Theil and Pietra coefficients as functions of the parameters are provided in Sections 3, 4, and 5, respectively. Proofs for the Gini coefficient and the Pietra index are given in an Appendix. In Section 6 we calculate inequality for South America as an illustrative example. Our note concludes with a remark in Section 7.

2. The Pareto-lognormal distribution

The density function for a random variable y that is distributed as Pareto-lognormal is

$$f(y)=\frac{\alpha}{y^{\alpha+1}}\exp\left\{\alpha\left(m+\frac{\alpha\sigma^2}{2}\right)\right\}\Phi\left(\frac{\ln y-m}{\sigma}-\alpha\sigma\right)$$

where (α,m,σ) are parameters and $\Phi(\cdot)$ denotes the standard normal distribution function. Its distribution function is given by

$$F(y)=\Phi\left(\frac{\ln y-m}{\sigma}\right)-\frac{1}{y^{\alpha}}\exp\left\{\alpha\left(m+\frac{\alpha\sigma^2}{2}\right)\right\}\Phi\left(\frac{\ln y-m}{\sigma}-\alpha\sigma\right)$$

The *k*-th moment for y is

$$E[y^k]=\frac{\alpha}{\alpha-k}\exp\left\{km+\frac{k^2\sigma^2}{2}\right\}\qquad \alpha>k$$

Now, suppose we have *n* Pareto-lognormal distributions indexed by *i* (and also later by *j*), and that the population share of the *i*-th country is λ_i. The density and distribution functions of their population-weighted mixture are given respectively by

$$f(y)=\sum_{i=1}^{n}\lambda_i f_i(y)=\sum_{i=1}^{n}\lambda_i\frac{\alpha_i}{y^{\alpha_i+1}}\exp\left\{\alpha_i\left(m_i+\frac{\alpha_i\sigma_i^2}{2}\right)\right\}\Phi\left(\frac{\ln y-m_i}{\sigma_i}-\alpha_i\sigma_i\right)$$

and

[1] See the review by Chotikapanich et. al (2018).

$$F(y)=\sum_{i=1}^{n}\lambda_i F_i(y)=\sum_{i=1}^{n}\lambda_i\left[\Phi\left(\frac{\ln y-m_i}{\sigma_i}\right)-\frac{1}{y^{\alpha_i}}\exp\left\{\alpha_i\left(m_i+\frac{\alpha_i\sigma_i^2}{2}\right)\right\}\Phi\left(\frac{\ln y-m_i}{\sigma_i}-\alpha_i\sigma_i\right)\right]$$

The mean of the mixture is $\mu=\sum_{i=1}^{n}\lambda_i\mu_i$ where μ_i is the mean of the *i*-th component. We assume throughout that $\alpha_i>1,$ so that the means exist for each of the component countries.

3. Gini coefficient

The Gini coefficient for the mixture is given by

$$G=-1+\frac{2}{\mu}\int_0^{\infty}y\,f(y)F(y)dy$$

$$=-1+\frac{2}{\mu}\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\lambda_j\left[\mu_i\Phi\left(h_{ij}\right)+e_{ij}c_{ij}\mu_j\Phi\left(d_{ij}\right)\right]$$

where

$$h_{ij}=\frac{m_i-m_j+\sigma_i^2}{\sqrt{\sigma_i^2+\sigma_j^2}}$$

$$e_{ij}=\left(\frac{\alpha_j-1}{(\alpha_i-1)(\alpha_i+\alpha_j-1)}\right)$$

$$c_{ij}=\exp\left\{\frac{\alpha_i^2}{2}\left(\sigma_i^2+\sigma_j^2\right)+\alpha_i\left(m_i-m_j-\sigma_j^2\right)\right\}$$

and

$$d_{ij}=\frac{m_j-m_i-\alpha_i\sigma_i^2-\left(\alpha_i-1\right)\sigma_j^2}{\sqrt{\sigma_i^2+\sigma_j^2}}$$

Proof: See the Appendix.

4. Theil Indices

Theil (1967) suggested two possible measures of inequality. Denoting them by T_0 and T_1, and using results in Sarabia et al. (2017), they are given by

$$T_0=\int_0^{\infty}\ln\left(\frac{\mu}{y}\right)f(y)dy=\ln\mu-\sum_{i=1}^{n}\lambda_i\int_0^{\infty}\ln y\,f_i(y)dy=\ln\mu-\sum_{i=1}^{n}\lambda_i\left(m_i+1/\alpha_i\right)$$

and

$$T_1 = \int_0^{\infty} \left(\frac{y}{\mu}\right) \ln\left(\frac{y}{\mu}\right) f(y)dy = \frac{1}{\mu}\sum_{i=1}^{n}\lambda_i \int_0^{\infty} y \ln y \, f_i(y)dy - \ln\mu$$

$$= \frac{1}{\mu}\sum_{i=1}^{n}\lambda_i \mu_i \left(m_i + \sigma_i^2 + \frac{1}{\alpha_i - 1} \right) - \ln\mu$$

5. Pietra Index

The Pietra index for the mixture is given by

$$P = \frac{1}{2\mu}\int_0^{\infty} |y-\mu| \, f(y)dy$$

$$= \sum_{i=1}^{n}\lambda_i \left[\Phi(e_i) - \frac{\mu_i}{\mu}\Phi(g_i) + \frac{1}{\mu^{\alpha_i}(\alpha_i - 1)} \exp\left\{ \alpha_i \left(m_i + \frac{\alpha_i \sigma_i^2}{2} \right) \right\} \Phi(h_i) \right]$$

where $e_i = (\ln\mu - m_i)/\sigma_i$, $g_i = (\ln\mu - m_i)/\sigma_i - \sigma_i$, and $h_i = (\ln\mu - m_i)/\sigma_i - \alpha_i\sigma_i$.

Proof: see the Appendix.

6. Example

To illustrate computations from the above results, in Table 1 we present estimates of inequality for South America and its component countries for the years 2001, 2006 and 2012. We have omitted Guyana, Suriname and French Guiana whose data were unavailable or incomplete. Their combined population is less than 0.4% of the total population of South America. Formulas for the inequality measures for the Pareto-lognormal components can be found in Hajargasht and Griffiths (2013). The parameter estimates are drawn from a subset of distributions estimated as a part of a larger project on the construction of a global panel of income distributions.[2] The data were expressed in terms of common currency units using the purchasing power parities from the UQICD website.[3] Also included in Table 1 are the population proportions that make up the weights in the mixture and annual mean incomes estimated from the parameters of the Pareto-lognormal distributions.

---

[2] The parameter estimates are available from the authors on request. The larger project is the Australian Reasearch Council Project DP140100673, Modelling Income Distributions over Space and Time. The panel of income distributions will be released later in 2020 on the UQICD website, uqicd.economics.uq.edu.au. The authors acknowledge the contributions of Prasda Rao and Dongjie Wu to the larger project.

[3] The purchasing power parities on the UQICD website cover the period 1970 to 2012; they are currently being extended to 2018.

Brazil is by far the largest country with half of the total population of South America. Uruguay and Chile have the highest per capita incomes. Examining the country-level inequality estimates from the alternative inequality measures, we find the measures broadly agree on the relative inequalities. For example, in 2006 all four measures point to Bolivia, Chile and Venezuela as the three countries with the highest level of inequality. In 2012 they agree on Brazil, Chile and Venezuela as having the highest level. In 2001, the Theil-1 measure gives Bolivia, Chile and Columbia as the three countries with the highest inequality, whereas Bolivia, Brazil and Columbia are the highest from the other three measures. With one exception, the direction of the change in inequality over the three years is consistent across all measures.[4] In all countries except for Venezuela, inequality was less in 2012 than it was in 2001. The decline from 2001 to 2012 was not monotonic for all countries, however. For Brazil and Paraguay, inequality increased from 2006 to 2012; for Peru and Uruguay, it increased from 2001 to 2006. Our calculations for South America as a whole reveal a decline in equality from 2001 to 2006 followed by a slight increase from 2006 to 2012.[5] Consistent with nine out of ten individual countries, inequality was lower in 2012 than it was in 2001.

7. Concluding Remark

We have derived formulas for inequality measures mixtures of Pareto-lognormal distributions and illustrated how they can be applied. The results provide a useful resource for researchers conducting large-scale projects for which the Pareto-lognormal distribution is utilised.

[4] The exception is Theil-0 for Paraguay between 2006 and 2012.

[5] The Theil-0 values in 2006 and 2012 are approximately the same.

Appendix

A1 Gini coefficient

To derive the Gini coefficient for the mixture of Pareto-lognormal distributions, we begin with the following lemma.

Lemma:

For $a < 0$ and $\sigma_1 > 0, \sigma_2 > 0$,

$$\int_0^\infty y^{a-1}\Phi\left(\frac{\ln y - b_1}{\sigma_1}\right)\Phi\left(\frac{\ln y - b_2}{\sigma_2}\right)dy = -\frac{1}{a}\left[\exp\left\{\frac{a^2\sigma_1^2}{2} + ab_1\right\}\Phi\left(\frac{b_1 - b_2 + a\sigma_1^2}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right) + \exp\left\{\frac{a^2\sigma_2^2}{2} + ab_2\right\}\Phi\left(\frac{b_2 - b_1 + a\sigma_2^2}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right)\right]$$

Proof

Let $u = (\ln y - b_1)/\sigma_1$, so that $\ln y = b_1 + \sigma_1 u$, $y = \exp\{b_1 + \sigma_1 u\}$, $dy = \sigma_1 \exp\{b_1 + \sigma_1 u\}$, and

$$\frac{\ln y - b_2}{\sigma_2} = \frac{\delta + \sigma_1 u}{\sigma_2}$$

where $\delta = b_1 - b_2$. Then,

$$\begin{aligned} I &= \int_0^\infty y^{a-1}\Phi\left(\frac{\ln y - b_1}{\sigma_1}\right)\Phi\left(\frac{\ln y - b_2}{\sigma_2}\right)dy \\ &= \sigma_1 \exp\{ab_1\}\int_{-\infty}^\infty \exp\{a\sigma_1 u\}\Phi(u)\,\Phi\left(\frac{\delta + \sigma_1 u}{\sigma_2}\right)du \\ &= \sigma_1 \exp\{ab_1\} I_0 \end{aligned}$$

To evaluate I_0, we use integration by parts. Let

$$X = \Phi(u)\,\Phi\left(\frac{\delta + \sigma_1 u}{\sigma_2}\right)$$

and $dW = \exp\{a\sigma_1 u\}$. Then,

$$I_0 = XW\Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} W\,dX$$

Consider

$$XW = \frac{\Phi(u)\Phi\left(\dfrac{\delta + \sigma_1 u}{\sigma_2}\right)}{\exp\{-a\sigma_1 u\}/a\sigma_1}$$

Given $a < 0$, it is clear that

$$\lim_{u \to \infty} XW = 0$$

For $u \to -\infty$, the limits of the numerator and denominator of XW both approach zero. Using l'Hôpital's rule, differentiating will lead to terms in the numerator which involve the normal density function; the term in u in the denominator will be unchanged. The normal density function contains terms such as $\exp\{-u^2/2\}$ which approach zero as $u \to -\infty$. Moreover, they approach zero at a faster rate than that for the term in the denominator. Thus,

$$\lim_{u\to-\infty} XW = 0$$

and

$$\begin{aligned} I_0 &= -\int_0^\infty W dX \\ &= -\frac{1}{a\sigma_1}\int_{-\infty}^{\infty} \exp\{a\sigma_1 u\}\left[\Phi\left(\frac{\delta+\sigma_1 u}{\sigma_2}\right)\phi(u) + \Phi(u)\phi\left(\frac{\delta+\sigma_1 u}{\sigma_2}\right)\frac{\sigma_1}{\sigma_2}\right]du \\ &= -\frac{1}{a\sigma_1}[A_1 + A_2] \end{aligned}$$

where

$$A_1 = \int_{-\infty}^{\infty} \exp\{a\sigma_1 u\}\Phi\left(\frac{\delta+\sigma_1 u}{\sigma_2}\right)\phi(u)\,du$$

$$A_2 = \int_{-\infty}^{\infty} \exp\{a\sigma_1 u\}\Phi(u)\phi\left(\frac{\delta+\sigma_1 u}{\sigma_2}\right)\frac{\sigma_1}{\sigma_2}du$$

Consider A_1, and focus on the terms in the exponent of $\exp\{a\sigma_1 u\}\phi(u)$, namely

$$a\sigma_1 u - \frac{u^2}{2} = -\frac{1}{2}(u - a\sigma_1)^2 + \frac{a\sigma_1^2}{2}$$

Then,

$$A_1 = \exp\left\{\frac{a^2\sigma_1^2}{2}\right\}\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{1}{2}(u-a\sigma_1)^2\right\}\Phi\left(\frac{\delta+\sigma_1 u}{\sigma_2}\right)du$$

Let $v = u - a\sigma_1$, so that $u = v + a\sigma_1$ and $du = dv$. Then,

$$\begin{aligned} A_1 &= \exp\left\{\frac{a^2\sigma_1^2}{2}\right\}\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{v^2}{2}\right\}\Phi\left(\frac{\delta + a\sigma_1^2 + \sigma_1 v}{\sigma_2}\right)dv \\ &= \exp\left\{\frac{a^2\sigma_1^2}{2}\right\}\Phi\left(\frac{\delta + a\sigma_1^2}{\sqrt{\sigma_1^2+\sigma_2^2}}\right) \end{aligned}$$

The last equality follows from the following result provided by Gupta and Pillai (1965).

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{z^2}{2}\right\}\Phi(cz+d)dz = \Phi\left(\frac{d}{\sqrt{1+c^2}}\right)$$

Now, consider A_2, and focus on the terms in the exponent of $\exp\{a\sigma_1 u\}\phi\left[(\delta+\sigma_1 u)/\sigma_2\right]$, namely

$$a\sigma_1 u - \frac{1}{2\sigma_2^2}(\delta + \sigma_1 u)^2 = -\frac{\sigma_1^2}{2\sigma_2^2}\left(u + \frac{\delta - a\sigma_2^2}{\sigma_1}\right)^2 + \frac{a^2\sigma_2^2}{2} - a\delta$$

Then,

$$A_2 = \exp\left\{\frac{a^2\sigma_2^2}{2} - a\delta\right\}\frac{\sigma_1}{\sigma_2}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{\sigma_1^2}{2\sigma_2^2}\left(u + \frac{\delta - a\sigma_2^2}{\sigma_1}\right)^2\right\}\Phi(u)du$$

Let

$$v = \frac{\sigma_1}{\sigma_2}\left(u + \frac{\delta - a\sigma_2^2}{\sigma_1}\right)$$

so that

$$u = \frac{\sigma_2}{\sigma_1}v - \frac{\delta - a\sigma_2^2}{\sigma_1}$$

and

$$du = \frac{\sigma_2}{\sigma_1}dv$$

Then,

$$A_2 = \exp\left\{\frac{a^2\sigma_2^2}{2} - a\delta\right\}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{v^2}{2}\right\}\Phi\left(\frac{-\delta + a\sigma_2^2 + \sigma_2 v}{\sigma_1}\right)dv$$

Using again the result from Gupta and Pillai (1965),

$$A_2 = \exp\left\{\frac{a^2\sigma_2^2}{2} - a\delta\right\}\Phi\left(\frac{-\delta + a\sigma_2^2}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right)$$

Collecting all the terms, we have

$$\begin{aligned}
I &= \sigma_1 \exp\{ab_1\} I_0 \\
&= \sigma_1 \exp\{ab_1\}\left[-\frac{1}{a\sigma_1}\left(A_1 + A_2\right)\right] \\
&= -\frac{1}{a}\left[\exp\left\{ab_1 + \frac{a^2\sigma_1^2}{2}\right\}\Phi\left(\frac{\delta + a\sigma_1^2}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right) + \exp\left\{ab_1 + \frac{a^2\sigma_2^2}{2} - a\delta\right\}\Phi\left(\frac{-\delta + a\sigma_2^2}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right)\right] \\
&= -\frac{1}{a}\left[\exp\left\{\frac{a^2\sigma_1^2}{2} + ab_1\right\}\Phi\left(\frac{b_1 - b_2 + a\sigma_1^2}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right) + \exp\left\{\frac{a^2\sigma_2^2}{2} + ab_2\right\}\Phi\left(\frac{b_2 - b_1 + a\sigma_2^2}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right)\right]
\end{aligned}$$

This completes the proof of the lemma. To use it to derive the Gini coefficient, we have

$$
\begin{aligned}
G &= -1+\frac{2}{\mu}\int_0^\infty yf(y)F(y)dy \\
&= -1+\frac{2}{\mu}\int_0^\infty y\sum_{i=1}^n \lambda_i f_i(y)\sum_{j=1}^n \lambda_j F_j(y)dy \\
&= -1+\frac{2}{\mu}\sum_{i=1}^n\sum_{j=1}^n \lambda_i\lambda_j\int_0^\infty yf_i(y)F_j(y)dy \\
&= -1+\frac{2}{\mu}\sum_{i=1}^n\sum_{j=1}^n \lambda_i\lambda_j\alpha_i \exp\left\{\alpha_i\left(m_i+\frac{\alpha_i\sigma_i^2}{2}\right)\right\}\left[I_1-\exp\left\{\alpha_j\left(m_j+\frac{\alpha_j\sigma_j^2}{2}\right)I_2\right\}\right]
\end{aligned}
$$

where

$$I_1=\int_0^\infty \frac{1}{y^{\alpha_i}}\Phi\left(\frac{\ln y-m_i-\alpha_i\sigma_i^2}{\sigma_i}\right)\Phi\left(\frac{\ln y-m_j}{\sigma_j}\right)dy$$

and

$$I_2=\int_0^\infty \frac{1}{y^{\alpha_i+\alpha_j}}\Phi\left(\frac{\ln y-m_i-\alpha_i\sigma_i^2}{\sigma_i}\right)\Phi\left(\frac{\ln y-m_j-\alpha_j\sigma_j^2}{\sigma_j}\right)dy$$

Applying the lemma to I_1 with $a=1-\alpha_i, b_1=m_i+\alpha_i\sigma_i^2$ and $b_2=m_j$ yields

$$
\begin{aligned}
I_1=\frac{1}{\alpha_i-1}\Bigg[&\exp\left\{m_i(1-\alpha_i)+\frac{\sigma_i^2}{2}\left(1-\alpha_i^2\right)\right\}\Phi\left(\frac{m_i-m_j+\sigma_i^2}{\sqrt{\sigma_i^2+\sigma_j^2}}\right) \\
&+\exp\left\{m_j(1-\alpha_i)+\frac{\sigma_j^2}{2}\left(1-\alpha_i\right)^2\right\}\Phi\left(\frac{m_j-m_i-\alpha_i\sigma_i^2+(1-\alpha_i)\sigma_j^2}{\sqrt{\sigma_i^2+\sigma_j^2}}\right)\Bigg]
\end{aligned}
$$

Applying the lemma to I_2 with $a=1-\alpha_i-\alpha_j, b_1=m_i+\alpha_i\sigma_i^2$ and $b_2=m_j+\alpha_j\sigma_j^2$ yields

$$
\begin{aligned}
I_2=\frac{1}{\alpha_i+\alpha_j-1}\Bigg[&\exp\left\{\frac{\sigma_i^2}{2}\left(1-2\alpha_j-\alpha_i^2+\alpha_j^2\right)+m_i(1-\alpha_i-\alpha_j)\right\}\Phi\left(\frac{m_i-m_j+(1-\alpha_j)\sigma_i^2-\alpha_j\sigma_j^2}{\sqrt{\sigma_i^2+\sigma_j^2}}\right) \\
&+\exp\left\{\frac{\sigma_j^2}{2}\left(1-2\alpha_i-\alpha_j^2+\alpha_i^2\right)+m_j(1-\alpha_i-\alpha_j)\right\}\Phi\left(\frac{m_j-m_i-\alpha_i\sigma_i^2+(1-\alpha_i)\sigma_j^2}{\sqrt{\sigma_i^2+\sigma_j^2}}\right)\Bigg]
\end{aligned}
$$

Substituting for I_1 and I_2 and simplifying gives

$$
\begin{aligned}
G=-1+\frac{2}{\mu}\sum_{i=1}^n\sum_{j=1}^n\lambda_i\lambda_j\Bigg[&\frac{\alpha_i}{\alpha_i-1}\left(\exp\left\{m_i+\frac{\sigma_i^2}{2}\right\}\Phi(h_{ij})+c_{ij}\exp\left\{m_j+\frac{\sigma_j^2}{2}\right\}\Phi(d_{ij})\right) \\
&-\frac{\alpha_i}{\alpha_i+\alpha_j-1}\left(c_{ji}\exp\left\{m_i+\frac{\sigma_i^2}{2}\right\}\Phi(d_{ji})+c_{ij}\exp\left\{m_j+\frac{\sigma_j^2}{2}\right\}\Phi(d_{ij})\right)\Bigg]
\end{aligned}
$$

where

$$h_{ij} = \frac{m_i - m_j + \sigma_i^2}{\sqrt{\sigma_i^2 + \sigma_j^2}}$$

$$c_{ij} = \exp\left\{\frac{\alpha_i^2}{2}\left(\sigma_i^2 + \sigma_j^2\right) + \alpha_i\left(m_i - m_j - \sigma_j^2\right)\right\}$$

and

$$d_{ij} = \frac{m_j - m_i - \alpha_i\sigma_i^2 - (\alpha_i - 1)\sigma_j^2}{\sqrt{\sigma_i^2 + \sigma_j^2}}$$

Recognizing that $\mu_i = \left[\alpha_i/(\alpha_i - 1)\right]\exp\left\{m_i + \sigma_i^2/2\right\}$, and simplifying, yields the final result

$$G = -1 + \frac{2}{\mu}\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\lambda_j\left[\mu_i\Phi(h_{ij}) + e_{ij}c_{ij}\mu_j\Phi(d_{ij})\right]$$

where

$$\begin{aligned} e_{ij} &= \left(\frac{\alpha_j - 1}{\alpha_j}\right)\left[\frac{\alpha_i}{\alpha_i - 1} - \frac{\alpha_j}{\alpha_i + \alpha_j - 1} - \frac{\alpha_i}{\alpha_i + \alpha_j - 1}\right] \\ &= \frac{\alpha_j - 1}{(\alpha_i - 1)(\alpha_i + \alpha_j - 1)} \end{aligned}$$

A2 Pietra index

The Pieta index is given by

$$\begin{aligned} P &= \frac{1}{2\mu}\int_0^{\infty}|\,y - \mu\,|\,f(y)dy \\ &= \int_0^{\mu} f(y)dy - \frac{1}{\mu}\int_0^{\mu} yf(y)dy \\ &= F(\mu) - F^{(1)}(\mu) \end{aligned}$$

where $F^{(1)}(y) = (1/\mu)\sum_{i=1}^{n}\lambda_i\mu_i F_i^{(1)}(y)$ is the first moment distribution. Using results in Hajargasht and Griffiths (2013),

$$F(\mu) = \sum_{i=1}^{n}\lambda_i\left[\Phi\left(\frac{\ln\mu - m_i}{\sigma_i} - \frac{1}{\mu^{\alpha_i}}\exp\left\{\alpha_i\left(m_i + \frac{\alpha_i\sigma_i^2}{2}\right)\right\}\Phi\left(\frac{\ln\mu - m_i}{\sigma_i} - \alpha_i\sigma_i\right)\right)\right]$$

and

$$F^{(1)}(\mu)=\frac{1}{\mu}\sum_{i=1}^{n}\lambda_i\mu_i\left[\Phi\left(\frac{\ln\mu-m_i}{\sigma_i}-\sigma_i\right)-\frac{1}{\mu^{\alpha_i-1}}\exp\left\{m_i(\alpha_i-1)+\frac{\sigma_i^2}{2}(\alpha_i^2-1)\right\}\Phi\left(\frac{\ln\mu-m_i}{\sigma_i}-\alpha_i\sigma_i\right)\right]$$

Subtracting the second term in $F(\mu)$ from the second term in $F^{(1)}(\mu)$, we have

$$\begin{aligned}
&\frac{1}{\mu}\sum_{i=1}^{n}\lambda_i\mu_i\left[\frac{1}{\mu^{\alpha_i-1}}\exp\left\{m_i(\alpha_i-1)+\frac{\sigma_i^2}{2}(\alpha_i^2-1)\right\}\Phi\left(\frac{\ln\mu-m_i}{\sigma_i}-\alpha_i\sigma_i\right)\right]\\
&\qquad-\sum_{i=1}^{n}\lambda_i\left[\frac{1}{\mu^{\alpha_i}}\exp\left\{\alpha_i\left(m_i+\frac{\alpha_i\sigma_i^2}{2}\right)\right\}\Phi\left(\frac{\ln\mu-m_i}{\sigma_i}-\alpha_i\sigma_i\right)\right]\\
&=\sum_{i=1}^{n}\lambda_i\left[\frac{1}{\mu^{\alpha_i}}\exp\left\{\alpha_i\left(m_i+\frac{\alpha_i\sigma_i^2}{2}\right)\right\}\Phi\left(\frac{\ln\mu-m_i}{\sigma_i}-\alpha_i\sigma_i\right)\left(\mu_i\exp\left\{-m_i-\frac{\sigma_i^2}{2}\right\}-1\right)\right]\\
&=\sum_{i=1}^{n}\lambda_i\left[\frac{1}{\mu^{\alpha_i}}\exp\left\{\alpha_i\left(m_i+\frac{\alpha_i\sigma_i^2}{2}\right)\right\}\Phi\left(\frac{\ln\mu-m_i}{\sigma_i}-\alpha_i\sigma_i\right)\left(\frac{1}{\alpha_i-1}\right)\right]
\end{aligned}$$

Thus,

$$\begin{aligned}
F(\mu)-F^{(1)}(\mu)=&\sum_{i=1}^{n}\lambda_i\Phi\left(\frac{\ln\mu-m_i}{\sigma_i}\right)-\frac{1}{\mu}\sum_{i=1}^{n}\lambda_i\mu_i\Phi\left(\frac{\ln\mu-m_i}{\sigma_i}-\sigma_i\right)\\
&+\sum_{i=1}^{n}\lambda_i\left[\frac{1}{\mu^{\alpha_i}(\alpha_i-1)}\exp\left\{\alpha_i\left(m_i+\frac{\alpha_i\sigma_i^2}{2}\right)\right\}\Phi\left(\frac{\ln\mu-m_i}{\sigma_i}-\alpha_i\sigma_i\right)\right]
\end{aligned}$$

References

Chotikapanich, D., Griffiths, W.E., Hajargasht, G. Karunarathne, W. D.S. Prasada Rao, D.S.: Using the GB2 income distribution. Econometrics 6(2), 21, 1-24 (2018)

Chotikapanich, D., Griffiths, W., Rao, D.S.P., Valencia, V.: Global income distributions and inequality, 1993 and 2000: incorporating country-level inequality modelled with beta distributions. Review of Economics and Statistics 94, 52-73 (2012)

Colombi, R.: A new model of income distribution: The Pareto-lognormal distribution. In: Dagum, C., Zenga, M. (eds.) Income and Wealth Distribution, Inequality and Poverty, pp.18-32. Springer, Berlin (1990)

Fabinger, M., Weyl E.G.: Functional forms for tractable economic models and the cost structure of international trade. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2194855 (2018)

Gupta, S., Pillai, S.: On linear functions of ordered correlated normal random variables. Biometrika 52, 367-379 (1965)

Giesen, K., Zimmermann, A., Suedekum, J.: The size distribution across all cities – double Pareto lognormal strikes. Journal of Urban Economics 68, 129-137 (2010)

Hajargasht, G., Griffiths, W.E.: Pareto-lognormal distributions: inequality, poverty, and estimation from grouped income data. Economic Modelling 33, 593-604 (2013)

Reed, W.J.: On the rank-size distribution for human settlements. Journal of Regional Science 42, 1-17 (2002)

Reed, W.J.: The Pareto law of incomes – an explanation and extension. Physica A 319, 469-486 (2003)

Reed, W.J., Jorgensen, M.: The double Pareto-lognormal distribution – a new parametric model for size distributions. Communications in Statistics Theory and Methods 33, 1733-1753 (2004)

Sarabia, J.M., Jordá, V., Remuzgo, L.: The Theil indices in parametric families of income distributions—a short review. Review of Income and Wealth 63, 867–80 (2017)

Theil, H.: Economics and Information Theory. North Holland, Amsterdam (1967)

Table 1

Inequality measures for South America and its component countries

	Argentina	Bolivia	Brazil	Chile	Colombia	Ecuador	Paraguay	Peru	Uruguay	Venezuela	Overall
Pop (%)											
2001	10.60	2.44	50.32	4.41	11.46	3.67	1.54	7.62	0.95	6.99	100
2006	10.49	2.51	50.18	4.38	11.53	3.77	1.58	7.50	0.89	7.18	100
2012	10.48	2.61	49.89	4.36	11.56	3.89	1.61	7.40	0.85	7.37	100
Mean											
2001	5790	1733	4281	5877	3984	3265	2805	3065	6497	4326	4287
2006	7291	2708	5363	9145	4934	4724	2785	4597	7750	5810	5546
2012	10770	3804	9582	15182	7369	6810	4716	7350	13555	11019	9333
Gini											
2001	0.4664	0.6103	0.5816	0.5724	0.5904	0.5384	0.5379	0.5158	0.4587	0.5222	0.5662
2006	0.4662	0.6061	0.5294	0.5528	0.5240	0.4894	0.4904	0.5175	0.4867	0.5401	0.5308
2012	0.4661	0.5232	0.5379	0.5410	0.5149	0.4687	0.4944	0.4614	0.4194	0.5946	0.5343
Theil 0											
2001	0.3818	0.7385	0.6538	0.5852	0.6787	0.5411	0.5398	0.4888	0.3651	0.5033	0.6180
2006	0.3814	0.7256	0.5198	0.5376	0.5011	0.4322	0.4344	0.4897	0.4126	0.5455	0.5254
2012	0.3814	0.5054	0.5274	0.5100	0.4869	0.3914	0.4334	0.3776	0.3048	0.6911	0.5252
Theil 1											
2001	0.4036	0.7421	0.6563	0.8594	0.6819	0.5431	0.5423	0.4904	0.4011	0.5049	0.6294
2006	0.4028	0.7286	0.5215	0.8063	0.5249	0.4340	0.4358	0.4999	0.4874	0.5468	0.5474
2012	0.4027	0.5073	0.5805	0.7828	0.4881	0.3929	0.4740	0.3790	0.3060	0.6939	0.5776
Pietra											
2001	0.3393	0.4569	0.4328	0.4227	0.4401	0.3973	0.3969	0.3791	0.3328	0.3843	0.4184
2006	0.3391	0.4533	0.3900	0.4069	0.3851	0.3582	0.3590	0.3802	0.3539	0.3987	0.3897
2012	0.3391	0.3850	0.3955	0.3977	0.3784	0.3420	0.3608	0.3363	0.3039	0.4436	0.3917