Statistical Methods & Applications (2001) 10: 41–48

SMA

Bayesian estimation of the Bonferroni index from a Pareto-type I population

G. M. Giorgi, M. Crescenzi

Dipartimento di Statistica, Probabilità e Statistiche Applicate. Università "La Sapienza", P.le Aldo Moro n. 5, 00185 Roma, Italy. (e-mail: {giovanni.giorgi,michele.crescenzi}@uniroma1.it)

Summary. The Bonferroni index (B) is a measure of income and wealth inequality, and it is particularly suitable for poverty studies. Since most income surveys are of a sample nature, we propose Bayes estimators of B from a Pareto/I population. The Bayesian estimators are obtained assuming a squared error loss function and, as prior distributions, the truncated Erlang density and the translated exponential one. Two different procedures are developed for a censored sample and for income data grouped in classes.

Key words: Bonferroni inequality index, Bayes estimator, Pareto/I distribution, truncated Erlang distribution, translated exponential distribution, squared error loss function.

1. Introduction

The Bonferroni (1930) index (B) was proposed as a descriptive measure of income inequality based on the comparison between partial means and the general mean (see Giorgi, 1998). Even though the B index is not well-known in statistical literature, it has the same properties as the famous Gini concentration ratio (R) and is also more sensitive to low levels of income distribution as stressed by some scholars (e.g. Giorgi and Mondani, 1995). This peculiarity has been studied in a recent paper in which B is used to construct a summary poverty measure (Giorgi and Crescenzi, 2001). For a more general review on the sensitivity of the Lorenz family's inequality measures, see Aaberge (2000).

Since the investigation on income distributions is mostly based on sample surveys, the study of some inferential aspects involved in the estimation of inequality measures is necessary. In this context Giorgi and Mondani (1994, 1995) have already derived the exact sampling distributions of the Bonferroni index, and their asymptotic properties, from a rectangular and exponential population.

In this paper B is estimated by following a Bayesian approach and by assuming that the population can be represented by the Pareto/I model. The Pareto distribution has been used in the statistical analysis of socio-economic phenomena since the end of XIX century and also over the last few years a number of authors (e.g. Arnold and Press, 1983, 1986; Moothathu, 1985, 1990; Ganguly *et al.* 1992) took it into consideration for making inference about inequality indices from income data.

When the Bayesian method is used, the choice of an appropriate prior distribution plays an important role. In the present paper we choose the truncated Erlang prior distribution since it is the natural conjugate for the Pareto/I model.

After a short review of the Bonferroni index and the Pareto/I distribution (Section 2), we show how the Bayes estimators of B may be obtained. There are many reasons, such as confidentiality, for which official agencies usually publish income data grouped in classes or in a censored form. Therefore we derive the Bayesian estimators of B with different procedures (Section 3 and Section 4) for these two cases. Finally, in Section 5, some suggestions on the method for assessing the hyperparameters are given.

2. The Bonferroni index in a Pareto/I population

We now introduce the Bonferroni inequality index and then we will examine its analytic expression in a Pareto/I model.

Let us suppose that the nonnegative and absolutely continuous random variable $X \in [0, +\infty)$ is the income and $F(x) = \int_0^x f(t)\,dt$ is its c.d.f. continuous and differentiable at least twice with $\mu = E(X) = \int_0^{+\infty} x\,f(x)\,dx \neq 0$ finite. The first incomplete moment and the partial mean of the probability distribution are

$$ {}_1F(x) = \frac{1}{\mu}\int_0^x t\,f(t)\,dt \tag{1} $$

$$ \mu_x = \frac{\int_0^x t\,f(t)\,dt}{\int_0^x f(t)\,dt} = \mu\left(\frac{{}_1F(x)}{F(x)}\right). \tag{2} $$

On the basis of the comparison between the partial means and the general mean, the *Bonferroni index B* is defined as

$$B = \frac{1}{\mu} \int_0^{+\infty} (\mu - \mu_x)\, dF(x) = 1 - \frac{1}{\mu} \int_0^{+\infty} \mu_x\, dF(x), \quad B \in [0,1]. \tag{3}$$

For computational purposes, it is important to show that B belongs to the class of income inequality linear measures introduced by Mehran (1976)

$$I = \frac{1}{\mu} \int_0^1 J(x)\, F^{-1}(x)\, dx \tag{4}$$

where J is a measurable real valued weighting function on $[0,1]$ and $F^{-1}(x) = \inf\{\, \xi : F(\xi) \geq x \,\}$.

The Bonferroni index can be derived (Nygård and Sandström, 1981, p. 210) by using the weighting function in (4), that is

$$J(x) = 1 + \ln(x). \tag{5}$$

For a suitable formalization it is better to consider a truncated r.v. X because, for tax purposes, the income receivers (*IR*s) below the threshold m may not be examined. We introduce the Pareto/I(m, θ) distribution for the absolutely continuous r.v. $X \in [m, +\infty)$, by its p.d.f.

$$f(x) = \frac{\theta\, m^{\theta}}{x^{\theta+1}}\, I\,(x > m) \tag{6}$$

where m, the scale parameter, and θ, the shape parameter, are positive. Since, in practice the minimum level of income m for paying tax is fixed, only θ is unknown.

The mean μ is finite only when $\theta > 1$ and, in this case, it is $\mu = m\theta/(\theta - 1)$. As the existence of μ is necessary for deriving the Bonferroni index, B can only be defined if $\theta > 1$. Under this condition, bearing in mind (4) and (5), the expression of B is

$$B = \Psi(2) - \Psi\left(1 + \frac{\theta - 1}{\theta}\right) \tag{7}$$

where the Digamma function $\Psi(z) = \Gamma'(z)/\Gamma(z)$ is the first derivative of the natural logarithm of $\Gamma(z) = \int_0^{+\infty} t^{z-1} e^{-t} dt$, $\forall z > 0$. If $n \in \mathrm{N}$, then $\Psi(n) = -\gamma + \sum_{k=1}^{n-1} \frac{1}{k}$, where

$$\gamma = \lim_{\ell \to +\infty} \left[\sum_{k=1}^{\ell} \frac{1}{k} - \ln(\ell) \right] \cong 0.57721567$$

is the Euler's constant.

Looking at (7), we can say that B is a strictly decreasing function of θ; this means that in a Pareto/I(m, θ) model, the inequality level is well-explained by the shape parameter θ. To be more precise, there is equidistribution when $\theta \to +\infty$, whereas if $\theta \to 1^+$ the inequality between *IR*s increases.

3. The case of a censored sample

In a Bayesian framework, besides considering the Pareto/I(m, θ) population with m known, we have to choose the *prior distribution* $g(\theta)$ for the r.v. θ. After having introduced $g(\theta)$ and constructed the *posterior distribution* $g(\theta \mid data)$, if we assume a squared error loss function, the Bayesian estimator of a generic function of θ, for instance $\varphi(\theta)$, will be $E[\varphi(\theta) \mid data]$. Therefore, to estimate the Bonferroni inequality index, it will be sufficient to remember that the function $\varphi(\theta)$, for B, is given by (7).

Therefore, let us suppose we have a sample, of dimension n, in which we can only observe $r < n$ outcomes of X, $(x_1, x_2, ..., x_r)$, corresponding to the *IR*s with an income not exceeding a fixed threshold w $(> m)$. On this point it should be noted that before sampling, r, contrary to n, is unknown. For instance, if w coincides with the poverty line (*PL*), the income distribution we obtain is the one theorized by Takayama (1979) and in this case B can be considered a poverty measure. We say the sample being studied is censored because there is a lack of information concerning the remaining $n - r$ *IR*s.

Note that if the sample had been complete (see Arnold and Press, 1983, p. 289) the likelihood function $L(\theta)$, in the Pareto/I(m, θ) model, would have been

$$L(\theta) = \theta^n m^{n\,\theta} \left(\prod_{i=1}^{n} x_i \right)^{-(\theta+1)} I\left(x_{(1)} > m\right) \tag{8}$$

where $x_{(1)} = \min\ \{x_1, x_2, ..., x_n\}$.

In the present case, it is suitable to introduce the statistics

$$P_w = w^{n-r} \left(\prod_{i=1}^{r} x_i \right) \quad ; \quad Z_w = \ln\ \left(m^{-n} P_w\right). \tag{9}$$

Seeing that $P(X > w) = 1 - F(w) = \left(\frac{m}{w}\right)^{\theta} = \lambda_w^{\theta}$, where $\lambda_w = \frac{m}{w}$, the likelihood function $L(\theta)$ for the censored sample is

$$L(\theta) = \frac{\theta^r m^{r\,\theta} \lambda_w^{(n-r)\,\theta}}{\left(\prod_{i=1}^{r} x_i \right)^{\theta+1}} \quad \propto \quad \theta^r e^{-\theta Z_w}, \quad \theta \in (\delta, +\infty). \tag{10}$$

Considering what we have just said in Section 2 about the existence of B in a Pareto/I(m, θ) population, we must take into account a truncated prior distribution since the r.v. θ is defined in $(\delta, +\infty)$ where the constant $\delta > 1$ is assumed as being known. In this paper we deal with the truncated Erlang prior distribution (TEPD) and one of its particular cases, that is the translated exponential.

3.1. Truncated Erlang prior distribution.

Under the assumption that the absolutely continuous r.v. $\theta \in (\delta, +\infty)$ has a TEPD($\beta, \ell\,; \delta$), its p.d.f. can be written as (see Ganguly *et al.* 1992, p. 96)

$$g(\theta) = \frac{\beta^{\ell}}{\Gamma(\ell, \delta\beta)}\theta^{\ell-1}e^{-\beta\,\theta}\ (\delta < \theta < +\infty, \delta > 1, \beta > 0, \ell = 1, 2, ...) \quad (11)$$

where $\Gamma(z, y) = \int_{y}^{+\infty} t^{z-1}\, e^{-t}\, dt, (y > 0)$. We derive the posterior distribution $g(\theta \mid data)$ following the Bayes theorem

$$g(\theta \mid data) = \frac{g(\theta) \cdot L(\theta)}{\int_{\Theta} g(\theta) \cdot L(\theta)\, d\theta} = \frac{g(\theta) \cdot L(\theta\,)}{\int_{\delta}^{+\infty} g(\theta) \cdot L(\theta\,)\, d\theta}. \quad (12)$$

Bearing in mind (10) and (11), the numerator in (12) becomes

$$g(\theta)\, L(\theta\,) \quad \propto \quad \frac{\beta^{\ell}}{\Gamma(\ell,\, \delta\beta)}\theta^{\ell+r-1}e^{-\,\theta\,(\beta+Z_w)} = \frac{\beta^{\ell}}{\Gamma(\ell, \delta\beta)}\theta^{\ell_*-1}e^{-\,\theta\beta_*} \quad (13)$$

where $\ell_* = \ell + r$ and $\beta_* = \beta + Z_w$.

Afterwards, considering that

$$\int_{\delta}^{+\infty} g(\theta)\, L(\theta\,)\, d\theta \quad \propto \quad \frac{\beta^{\ell}}{\Gamma(\ell, \delta\beta)} \int_{\delta}^{+\infty} \theta^{\ell_*-1}e^{-\,\theta\beta_*}\, d\theta =$$

$$\frac{\beta^{\ell}}{\Gamma(\ell, \delta\beta)} \cdot \frac{\Gamma(\ell_*, \delta\beta_*)}{\beta_*^{\ell_*}} \quad (14)$$

the posterior distribution can be expressed by

$$g(\theta \,\Big|\, data) = \frac{\beta_*^{\ell_*}}{\Gamma(\ell_*, \delta\,\beta_*)}\theta^{\ell_*-1}e^{-\,\theta\beta_*} = TEPD(\beta_*, \ell_*; \delta)\ . \quad (15)$$

It is well-known that, in a Bayesian and decisional framework, a loss function $l(T, \theta)$ must be defined for evaluating the error we make using the statistic $T = h(X_1, X_2, ..., X_n)$ to estimate the parameter $\theta, \forall\theta \in \Theta$. As just said, we shall adopt the squared error loss function $l(T, \theta) = (T - \theta)^2$. That choice allows us to identify the Bayesian estimator of a generic function $\varphi\,(\theta)$ with its mean $E\,[\,\varphi\,(\theta) \mid data\,]$. So, by combining (7) with (15), the Bayes estimator of the Bonferroni index for a censored sample is

$$\hat{B}_1 = \frac{\beta_*^{\ell_*}}{\Gamma(\ell_*, \delta\beta_*)} \int_{\delta}^{+\infty} \left[\Psi(2) - \Psi\left(1 + \frac{\theta - 1}{\theta}\right)\right] \theta^{\ell_*-1}e^{-\,\theta\beta_*}\, d\theta. \quad (16)$$

If we put $\ell = 1$ in (11), we obtain as particular case the translated exponential density (see Bhattacharya *et al.* 1999, p. 250)

$$g(\theta) = \beta\, e^{-\beta\,(\theta-\delta)}, (\delta < \theta < +\infty, \delta > 1, \beta > 0). \qquad (17)$$

Hence, by substituting $\ell = 1$ in (16), the Bayesian estimator of B becomes

$$\hat{B}_2 = \frac{(\beta_*)^{r+1}}{\Gamma(r+1, \delta\beta_*)} \int_{\delta}^{+\infty} \left[\Psi(2) - \Psi\left(1 + \frac{\theta - 1}{\theta}\right)\right] \theta^r e^{-\theta\beta_*}\, d\theta. \qquad (18)$$

4. Bayes estimator of B for grouped data

In the statistical analysis of income distribution, as discussed in the introduction, available data are often grouped in classes therefore suggesting that proper estimators for this kind of representation should be proposed.

If the assumptions in Section 2 for the r.v. X hold, let us consider the case in which we are not able to observe the specific value of X_i, $i = 1, ..., n$, and we only know whether or not it belongs to the income class $A_j = (a_{j-1}, a_j]$, $j = 1, ..., k$, with $a_0 = m$ and $a_k = +\infty$. Obviously the frequencies n_j of the observations that fall in the interval A_j, $j = 1, ..., k$ and $n = \sum_{j=1}^{k} n_j$, are known.

When sample data are grouped in k income classes, the likelihood function $L(\theta)$ can be computed bearing in mind that

$$P(X > a_j) = 1 - F(a_j) = \left(\frac{m}{a_j}\right)^{\theta} = \lambda_j^{\theta} \qquad (19)$$

where $\lambda_j = \frac{m}{a_j}$, $j = 1, ..., k-1$ and $\lambda_0 = 1$.

Therefore

$$L(\theta) \quad \propto \quad \left(1 - \lambda_1^{\theta}\right)^{n_1} \left(\lambda_1^{\theta} - \lambda_2^{\theta}\right)^{n_2} \cdot ... \cdot \left(\lambda_{k-2}^{\theta} - \lambda_{k-1}^{\theta}\right)^{n_{k-1}} \left(\lambda_{k-1}^{\theta}\right)^{n_k}. \qquad (20)$$

Using the binomial expansions in (20), we can write (see Ganguly *et al.* 1992)

$$L(\theta) \quad \propto \quad \sum\nolimits^{*} e^{-\xi\theta}, \quad \theta \in (\delta, +\infty) \qquad (21)$$

where

$$-\xi \equiv -\xi(i_1, i_2, ..., i_{k-1}) =$$
$$\left\{\sum_{s=0}^{k-2} (n_s - i_s + i_{s+1}) \ln \lambda_s\right\} + (n_{k-1} - i_{k-1} + n_k) \ln \lambda_{k-1},$$

and $\sum^{*} \equiv \sum_{i_1=0}^{n_1} \sum_{i_2=0}^{n_2} ... \sum_{i_{k-1}=0}^{n_{k-1}} \left\{\prod_{j=1}^{k-1} \binom{n_j}{i_j}\right\} (-1)^{\sum_{j=1}^{k-1}(n_j - i_j)}$.

Making use of the prior TEPD$(\beta, \ell; \delta)$ seen in Section 3.1, the posterior distribution for the case being studied is

$$g(\theta \mid data) = \frac{\sum^{*} \theta^{\ell-1} e^{-\theta\beta'}}{\sum^{*} (\beta')^{-\ell} \Gamma(\ell, \delta\beta')}, \quad \theta \in (\delta, +\infty) \tag{22}$$

where $\beta' \equiv \beta'(i_1, i_2, ..., i_{k-1}) = \beta + \xi\,(i_1, i_2, ..., i_{k-1}) = \beta + \xi$. Still having a squared error loss function, the Bayesian estimator of the Bonferroni index, for a sample grouped in classes, is

$$\hat{\hat{B}}_1 = \frac{\sum^{*} \int_{\delta}^{+\infty} \left[\Psi(2) - \Psi\left(1 + \frac{\theta-1}{\theta}\right)\right] \theta^{\ell-1} e^{-\theta\beta'} d\theta}{\sum^{*} (\beta')^{-\ell} \Gamma(\ell, \delta\beta')}. \tag{23}$$

When $g(\theta)$ has the translated exponential form, the Bayes estimator of B becomes

$$\hat{\hat{B}}_2 = \frac{\sum^{*} \int_{\delta}^{+\infty} \left[\Psi(2) - \Psi\left(1 + \frac{\theta-1}{\theta}\right)\right] e^{-\theta\beta'} d\theta}{\sum^{*} (\beta')^{-1} e^{-\delta\beta'}}. \tag{24}$$

By using the expressions (16) and (23), we have obtained the Bayesian estimators of B from a Pareto/I(m, θ) population in two different cases in which the income sampling information is available in an incomplete form. The circumstance of a censored distribution arises when data come from tax-returns and, to avoid the problems linked to the weak reliability of high incomes, it is used to truncate the distribution above the point w. On the other hand, the representation in classes is due to the fact that the official agencies collecting income data are obliged to keep the information confidential. The choice of a truncated prior distribution depends on the fact that the condition for the existence of B in the Pareto/I(m, θ) model, demands a $(\delta, +\infty)$ range with $\delta > 1$ for the shape parameter θ (see Section 3).

5. A remark on the assessment of hyperparameters

In this paper we have used (11) and (17) as priors for the Pareto-type I shape parameter θ. To fix a single prior distribution, we may assess the values of hyperparameters using prior information on the income distribution. Since many surveys on personal incomes provide the group means, on the basis of past data we can specify a set of expected group means for the study at hand. From these means we obtain an empirical distribution for θ bearing in mind that, in a Pareto/I population where the scale parameter m is known and $\theta > 1$, the mean is $m\theta/(\theta - 1)$. By comparing the percentiles of this distribution with the percentiles derived from the assumed prior [(11) or (17)], the hyperparameters can be assessed with *ad hoc* algorithms that for the truncated Erlang distribution and the translated exponential one are developed in Ganguly *et al.* (1992, pp. 102-104) and Bhattacharya *et al.* (1999, p. 256), respectively.

References

Aaberge R (2000) Characterizations of Lorenz curves and income distributions, Social Choice and Welfare **17**(4): 639–653

Arnold BC, Press SJ (1983) Bayesian inference for Pareto populations, Journal of Econometrics **21**: 287–306

Arnold BC, Press SJ (1986) Bayesian analysis of censored or grouped data from Pareto populations. In: Bayesian Inference and Decision Techniques, Essays in Honor of Bruno de Finetti. (Goel PK, Zellner A, eds.) North-Holland, Amsterdam 157–173

Bernardo JM, Smith AFM (1994) Bayesian Theory. Wiley, Chichester

Bhattacharya SK, Chathurvedi A, Sing NK (1999) Bayesian estimation for the Pareto Income distribution. Statistical Papers **40**(3): 247–262

Bonferroni CE (1930) Elementi di Statistica Generale. Seeber. Firenze

Ganguly A, Sing NK, Chouduri H, Bhattacharya SK (1992). Bayesian Estimation of the Gini Index for the PID, Test **1**(1): 93–104

Giorgi GM (1998) Concentration Index, Bonferroni, Encyclopedia of Statistical Sciences, Update, 2, Wiley, NY 141–146

Giorgi GM, Crescenzi M (2001) A proposal of poverty measures based on the Bonferroni inequality index, Metron, **59** (3–4): 3–16

Giorgi GM, Mondani R (1994) The exact sampling distribution of the Bonferroni concentration index, Metron, **52** (3–4): 5–41

Giorgi GM, Mondani R (1995) Sampling distribution of the Bonferroni inequality index from Exponential population, Sankhya, Series B **57** (1): 10–18

Mehran F (1976) Linear measures of income inequality, Econometrica **44** (4): 805–809

Moothathu TSK (1985) Sampling distributions of the Lorenz Curve and Gini Index of the Pareto Distribution, Sankhya, Series B, **47** (2): 247–258

Moothathu TSK (1990) The best estimator and a strongly consistent asymptotically normal unbiased estimator of Lorenz curve, Gini index and Theil entropy index of Pareto distribution, Sankhya, Series B **52** (1): 115–127

Nygård F, Sandström A (1981). Measuring Income Inequality. Stockolm: Almqvist & Wiksell International

Takayama N (1979) Poverty, income inequality, and their measures: Professor Sen's axiomatic approach reconsidered, Econometrica **47** (3): 747–759