Exact Distribution of the Gini Concentration Index from a Skew Normal Sample

Corrado Crocetta
Dipartimento di Scienze Economiche, Matematiche e Statistiche, Università di Foggia
Via IV Novembre 1, - 71100 Foggia, e-mail: c.crocetta@unifg.it

Nicola Loperfido
Istituto di Scienze Economiche, Università degli Studi di Urbino
Via Saffi, 42 – 61029 (PU) e-mail: nicola@econ.uniurb.it

Summary: In this paper we demonstrate that the exact sampling distribution of the Gini Concentration Index from a Skew Normal population, is an Extended Skew Normal distribution. This result can be easily applied to the particular case of samples from a Normal population.

Keywords: Skew Normal, L-statistics, Small Sample, Gini Concentration Ratio, Exact Distribution.

1. Introduction

The Gini concentration ratio

\[ R = \frac{\sum_{i=1}^{n} (2i - n - 1) x_{(i)}}{(n-1) \sum_{i=1}^{n} x_{(i)}} \]

has been extensively used in the study of distribution inequality.

A major statistical limitation of \( R \) is the absence and the intractability of appropriate sampling distribution (Hart, 1971, Nygard Sandström 1981 p. 372).

Hoeffding (1948) demonstrated that, under some regularity conditions, \( R \) has a normal asymptotical distribution. Yet, only few results are available about the exact distributions of \( R \).

A summary of the results about the asymptotic and exact distributions of \( R \) is in Nygard Sandström (1981) and Girone (1987), while Giorgi (1992, 1999, 2003), and Tarsitano (2004) are suggested for a more detailed review.

In the next section there are some general results about the Skew Normal distribution and one of its extensions. Section 3 states that the exact distribution of the Gini index from a Skew Normal distribution is related to an Extended Skew Normal distribution, whereas Section 4 proves the main results stated in section 3.
2. The Skew Normal distribution and one of its extensions.

The Skew Normal (hereafter SN) class of densities independently appeared several times in statistical literature [Roberts (1966), O'Hagan and Leonard (1976), Aigner and Lovell (1977)]. The present name was given by Azzalini (1985) and it was generalized to the multivariate case by Azzalini and Della Valle (1996) and Azzalini and Capitanio (1999).

The SN class of densities extends the Normal model by allowing a shape parameter to account for skewness. The density function of the generic element of the class is

$$f(x; \mu, \psi, \lambda) = \frac{2}{\psi} \phi\left(\frac{x - \mu}{\psi}\right) \Phi\left(\frac{x - \mu}{\psi}\right) \quad x, \lambda, \mu \in R; \quad \psi = R^+ \quad (1)$$

where \(\phi(\cdot)\) and \(\Phi(\cdot)\) denote the standard normal probability density function (pdf) and cumulative distribution function (cdf), respectively. \(X \sim SN(\mu, \psi, \lambda)\) shall be written to denote a random variable with density (1). The parameter \(\lambda\) controls skewness, which is positive when \(\lambda > 0\) and negative when \(\lambda < 0\). When \(\lambda = 0\) we have \(X \sim N(\mu, \psi)\). Despite skewness, these distributions resemble the normal ones in several ways: they are unimodal, their support is the real line and the square of a variable, whose distribution is skew-normal, follows a chi-square distribution.

Figure 1 shows some examples of Skew Normal densities with location parameter 0, scale parameter 1 and shape parameters 0, \(\infty\), -4 and 4 respectively.

**Figure 1** Examples of Skew Normal Densities with different values of \(\lambda\).
This generalization of the normal law has the advantage of being mathematically tractable and easily interpretable (Azzalini, 1985). It is more flexible and adaptable to real data, in particular, in the rather frequent case of phenomena whose empirical outcomes that behave in a non-normal way, yet still retaining some similarity with the normal distribution.

For the standard case $X \sim SN(0, 1, \lambda)$ where $\lambda = \delta / \sqrt{1 - \delta^2}$ the expectation, the variance, the skewness and the kurtosis index are respectively:

$$E(X) = \frac{\sqrt{2}}{\pi} \cdot \delta,$$
$$Var(X) = 1 - \frac{2}{\pi} \cdot \delta^2,$$
$$Sk = \frac{\sqrt{2/\pi} \cdot \delta^3 (4/\pi - 1)}{\sqrt{Var(X)^3}},$$
$$K = \frac{3\left(1 - \frac{2}{\pi} \cdot \delta^2 \right)^2 + 8 \cdot \delta^4 \left(1 - \frac{3}{\pi}\right)}{Var^2(X)}$$

The Skew Normal distribution can be obtained by a transformation method: the maximum of a bivariate normal distribution is a Skew Normal distribution; or by a conditioning method: the conditional distribution of a bivariate normal is a Skew Normal.

Hereafter, we give the definition of a new multivariate distribution. This is called Extended Skew Normal because it can be seen as a generalization of the Skew Normal distribution.

**EXTENDED SKEW NORMAL DISTRIBUTION (Arellano-Valle, Genton, 2003)**

A random vector $X \in \mathbb{R}^p$ is said to have an Extended Skew Normal distribution $ESN(\xi, \Psi, \Lambda, \Gamma)$ if its pdf is

$$f(x, \xi, \Psi, \Lambda, \Gamma) = \frac{\phi_p\left(x - \xi | \Psi \right)}{\Phi_p\left(0 | \Gamma \right)} \cdot \left[\Lambda \Psi^{-1} (x - \xi) | \Gamma - \Lambda \Psi^{-1} \Lambda \right]$$

where $\phi( \cdot | \Sigma)$ and $\Phi( \cdot | \Sigma)$ are the pdf and the cdf of a Normal distribution $N(0; \Sigma)$.

As it can be seen in section 3, the sampling distribution of the Gini Concentration Index from a Skew Normal population coincides with an Extended Skew Normal distribution.

### 3. The exact sampling distribution of R

We shall denote with $\Phi(x; \Sigma)$ the cdf of $N(\theta; \Sigma)$ evaluated in $x$. Moreover, we shall write $\Delta$ to denote the difference matrix of order $(n-1) \times n$ of the form

$$\Delta = \begin{pmatrix}
-1 & 1 & 0 & \ldots & 0 \\
0 & -1 & 1 & \ldots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \ldots & -1
\end{pmatrix}$$
THEOREM 1: Let \( \{X_1, \ldots, X_n\} \) be a random sample from \( SN(\mu; \psi; \lambda) \) and let \( R \) be the corresponding Gini statistic. Then

\[
P(R \leq r) = n! (\Phi(b|\Xi_1) + \Phi(-b|\Xi_2))
\]

The first component of the vector \( b \) is \( r n \mu (n-1) \), the second component is \( n \mu \) and all the remaining components are zero. The block matrices \( \Xi_1, \Xi_2 \) are defined as follows:

\[
\Xi_1 = \begin{pmatrix} a' & -a' & a' & 0 & \tilde{\delta} \\
-1' & n & 0 & -\delta' & 0 \\
\Delta & 0 & \Delta & \delta & 0 \\
\tilde{\delta} & -\delta & \delta & I & 0 \\
\end{pmatrix} \quad \Xi_2 = \begin{pmatrix} a' & -a' & a' & 0 & \tilde{\delta} \\
-1' & n & 0 & \delta' & 0 \\
-\Delta & 0 & \Delta & \delta & 0 \\
-\tilde{\delta} & \delta & \delta & I & 0 \\
\end{pmatrix}
\]

where \( \lambda = \delta / \sqrt{1 - \delta^2} \) and the \( i \)-th component of the vector \( a \) is \( 2(i-1)+(r+1)(1-n) \).

The advantage of representing \( P(R \leq r) \) in the above form becomes apparent when considering the fundamental results by Miwa, Hayter and Kuriki (2003), which lead to an algorithm for the accurate evaluation of \( \Phi(x|\Sigma) \).

The results obtained for the Gini Concentration ratio can be easily generalized to any ratio of L-statistics. This generalization can be applied to find the exact distribution of a wide class of concentration measures called Gini Family by Nygard and Sandström.

If we remember that the Normal distribution is a particular case of the Skew Normal with \( \lambda = 0 \) theorem 1 can be used to find the exact distribution of the Gini Index for samples from a Normal population.

In order to prove theorem 1 we have to remember some previous results.

THEOREM 2 (Crocetta, Loperfido, 2003): Let \( l = (L_1, \ldots, L_p)' = \Omega y \) be a linear transformation of the vector \( y = (Y_1 \leq \ldots \leq Y_n) \), whose elements are the order statistics \( Y_1 \leq \ldots \leq Y_n \) corresponding to the random sample \( X_1, \ldots, X_n \) from \( SN(\lambda) \). Then \( l \sim ESN(0; \Omega \Omega'; \Lambda; \Gamma) \), where:

\[
\Lambda = \begin{pmatrix} \Lambda \Omega' \\
\delta \Omega \\
\end{pmatrix} \quad \Gamma = \begin{pmatrix} \Delta & \delta \Delta \\
\delta \Delta & I \\
\end{pmatrix} \quad \delta = \frac{\lambda}{\sqrt{1 + \lambda^2}}
\]

THEOREM 3: The cdf of \( X \sim ESN(\theta; \Psi; \Lambda; \Gamma) \) is:

\[
F_X(a) = \frac{\Phi(a; \theta | \Xi)}{\Phi(\theta | \Gamma)}
\]

\[
\Xi = \begin{pmatrix} \Psi & \Lambda \\
\Lambda & \Gamma \\
\end{pmatrix}
\]

where \( \Psi = \Omega \Omega' \).
4. Proof of Theorems

In order to prove theorem 1, the results of theorems 2 and 3 are needed. Theorem 3 is to be the first to prove.

PROOF OF THEOREM 3: Let $Z$ and $W$ be two random vectors whose joint distribution is normal, with $\theta$ mean and covariance matrix $\Xi$:

$$
\begin{pmatrix}
Z \\
W
\end{pmatrix} \sim N(\theta; \Xi)
$$

By definition, $\Phi(\cdot ; \Sigma)$ is the cdf of $N(\theta; \Sigma)$. Hence:

$$
P(Z \leq a; W \leq 0) = \Phi(a; \theta | \Xi) \quad P(W \leq 0) = \Phi(0; \Gamma)
$$

a standard application of Bayes’ theorem leads to the density of $Z|W \leq 0$:

$$
f_{Z|W \leq 0}(a) = \frac{\phi(a | \Psi)}{\Phi(\theta | \Gamma)} \cdot \Phi\left[\Lambda^{-1}a | \Gamma - \Lambda \Psi^{-1} \Lambda^T\right]
$$

where $\phi(\cdot ; \Xi)$ is the density of $N(\theta; \Sigma)$. By definition, the above equation denotes the density of $ESN(\theta; \Psi; \Lambda; \Gamma)$. Hence:

$$
F_X(a) = \frac{P(Z \leq a; \ W \leq a)}{P(W \leq 0)} = \frac{\Phi(a; \theta | \Xi)}{\Phi(\theta | \Gamma)}
$$

and the proof is complete.

PROOF OF THEOREM 1: The Gini statistic is scale invariant. Thus, we can assume that $\psi = 1$ without loss of generality. Let $y = (Y_1, \ldots, Y_n)$ be the vector of order statistics corresponding to $X_1, \ldots, X_n$ and recall the definition of $R$: 

$$
R = \frac{\sum (2i - n - 1)Y_i}{(n - 1)\sum Y_i}
$$

Apply now the additive law:

$$
P(R \leq r) = F(r) = P\left[\sum \frac{(2i - n - 1)Y_i}{(n - 1)\sum Y_i} \leq r; \sum Y_i > 0\right] + P\left[\sum \frac{(2i - n - 1)Y_i}{(n - 1)\sum Y_i} \leq r; \sum Y_i \leq 0\right]
$$

A little algebra leads to the following:

$$
F(r) = P\left[\sum (2i - n - 1 - rn + r)Y_i \leq 0; \sum Y_i > 0\right] + P\left[\sum (2i - n - 1 - rn + r)Y_i \geq 0; \sum Y_i \leq 0\right]
$$
The definition of $a$ implies that
\[
F(r) = P[a'y \leq 0; I'y > 0] + P[a'y \geq 0; I'y \leq 0] = P[a'y \leq 0; -I'y < 0] + P[-a'y \leq 0; I'y \leq 0]
\] (3)

By definition, $a'y$ and $I'y$ are L-statistics from a Skew Normal. Hence theorem 2 holds and
\[
\begin{pmatrix}
a'y \\
-1'y
\end{pmatrix} \sim \text{ESN}\left[\begin{pmatrix} r\mu(n-1) \\
-n\mu
\end{pmatrix}; \begin{pmatrix} a'a & -1'a \\
-a'1 & n
\end{pmatrix}; \begin{pmatrix} \Delta a & 0 \\
\delta a & -\delta1
\end{pmatrix}; \begin{pmatrix} \Delta\Delta' & \delta\Delta \\
\delta\Delta' & I
\end{pmatrix}\right]
\]
\[
\begin{pmatrix}
-a'y \\
1'y
\end{pmatrix} \sim \text{ESN}\left[\begin{pmatrix} r\mu(n-1) \\
n\mu
\end{pmatrix}; \begin{pmatrix} a'a & -1'a \\
-a'1 & n
\end{pmatrix}; \begin{pmatrix} -\Delta a & 0 \\
-\delta a & \delta1
\end{pmatrix}; \begin{pmatrix} \Delta\Delta' & \delta\Delta \\
\delta\Delta' & I
\end{pmatrix}\right]
\]

Theorem 3 implies that:
\[
P[a'y \leq 0; -I'y < 0] = \frac{\Phi[r\mu(n-1); n\mu; \theta|\Xi_1]}{\Phi(\theta|\Gamma)}
\]
\[
P[-a'y \leq 0; -I'y \leq 0] = \frac{\Phi[r\mu(1-n); -n\mu; \theta|\Xi_2]}{\Phi(\theta|\Gamma)}
\]

where $\Gamma$ has the same meaning as in theorem 2. By assumption, $X_1, \ldots, X_n$ are i.i.d., so that
\[
\Phi(\theta|\Gamma) = P(X_1 \leq \ldots \leq X_n) = \frac{1}{n!}
\]

Recall now equation (3):
\[
P(R \leq r) = n!\left[\Phi(r\mu(n-1); n\mu; \theta|\Xi_1) + \Phi(r\mu(1-n); -n\mu; \theta|\Xi_2)\right]
\]

and the proof is complete.

References


