

This is the authors' accepted manuscript for the article: Mehdi and Stengos (2014) Empirical likelihood-based inference for the generalized entropy class of inequality measures. *Economics Letters* 123: 54 - 57. The final publication is available from: <https://dx.doi.org/10.1016/j.econlet.2014.01.015>

# Empirical likelihood-based inference for the generalized entropy class of inequality measures

Tahsin Mehdi, Thanasis Stengos\*

\* Department of Economics and Finance, University of Guelph, Guelph, ON N1G 2W1, Canada. Tel.: +1 519 824 4120x53917; fax: +1 519 763 8497. E-mail: tstengos@uoguelph.ca

October 28, 2013

## Abstract

We propose an empirical likelihood-based method of inference for comparing inequality between two populations. A series of Monte Carlo experiments are used to assess our method's finite sample performance. We illustrate our approach using some Canadian household income data.

*JEL Classification: C1, C12, I3*

*Keywords: Empirical likelihood, Inequality measures*

# 1 Introduction

Ever since the work of Atkinson (1970), there has been significant research interest in economic inequality and poverty. Although the measurement of inequality and poverty are important, statistical inference for such measures have gained considerable interest in recent years. The work of Kakwani (1993), Zheng (2001), Biewen (2002) and Davidson and Flachaire (2007) serve to highlight the importance of statistical inference in measuring inequality and poverty rather than just the incidence.

The growing body of literature surrounding the theory of inequality measurement has been accompanied by increasing availability of income data distribution which have armed researchers with the capability to conduct more sophisticated analyses. Statistical inference for inequality measures was largely neglected until the work of Cowell (1989). Recently, Thompson (2010) derived the asymptotic properties of vector measures of inequality (and poverty). He argued that since there is often no “best” measure of inequality or poverty, multiple measures could be used.

Our method of inference relies on empirical likelihood (EL), a powerful nonparametric statistical method pioneered by Owen (1988, 1990). An advantage of empirical likelihood is that no assumptions are needed regarding the underlying distribution of the data. Thompson (2013) used the approach for making inference on poverty measures which utilize relative poverty lines. His main focus was to compare poverty between two subgroups of a population that share a common poverty line. We depart from focusing on poverty measures and turn our attention to inequality measures (more specifically, we limit our focus to the generalized entropy class of inequality measures).

The remainder of this paper is organized as follows. In Section 2, we provide a brief overview of inequality measures. In Section 3, we present our methodology. In Section 4, we examine the finite sample performance of our method using a Monte Carlo simulation. In Section 5, we demonstrate the practicality of our method using an empirical application.

## 2 Inequality measures

In this section, we provide a basic overview of the measurement of inequality. For a more thorough treatment of the literature, see Cowell (2011) or Cowell (2000). Before proceeding, we need to introduce some notation. Following Thompson (2010), we generalize our approach for vector measures of inequality. Let  $Y = (Y_1, \dots, Y_J)'$  be a random vector whose value is determined by a set of attributes (e.g., income, education, etc.) for an individual from a certain population. In the case where we are interested in only one attribute but we want to consider  $J$  distinct measures, we will have  $Y_j = Y_k$  for all  $j, k$ . Let  $F_j$  be the distribution function of  $Y_j$ .

There are several different scalar measures of inequality that exist in the literature. We focus exclusively on the generalized entropy class of measures which fulfill the most widely accepted axioms including decomposability (see, e.g., Cowell, 2000).<sup>1</sup> For the random vector  $Y_j$ , such measures can be written as  $I_j = E_j(h_j(y_j, \mu_j, \alpha_j))$  where  $E_j$  denotes expectation under distribution  $F_j$ ,  $h_j(y_j, \mu_j, \alpha_j)$  is some real-valued function,  $\mu_j$  is the mean of  $F_j$ , and  $\alpha_j$  is an exogenous parameter (and thus its choice is subjective). Formally, we have

$$I_j = \int h_j(y_j, \mu_j, \alpha_j) dF_j(y_j),$$

---

<sup>1</sup>The Atkinson class of inequality measures, and the Gini index are some of the other well established measures of inequality.

where  $\mu_j = \int y_j dF_j(y_j)$  and

$$h_j(y_j, \mu_j, \alpha_j) = \begin{cases} [(y_j/\mu_j)^{\alpha_j} - 1]/(\alpha_j^2 - \alpha_j) & \alpha_j \neq 0, 1 \\ -\log(y_j/\mu_j) & \alpha_j = 0 \\ y_j \log(y_j/\mu_j)/\mu_j & \alpha_j = 1. \end{cases}$$

Let  $\mu = (\mu_1, \dots, \mu_J)'$  be the vector of means. A vector of inequality measures can be written as  $I = (I_1, \dots, I_J)'$ .

### 3 Empirical likelihood-based inference

The empirical likelihood method was first brought to the forefront by Owen (1988, 1990). It is a nonparametric method of inference and an alternative to the bootstrap. For an extensive overview, see Owen (2001).

The basic framework can be explained as follows. Let  $y_1, \dots, y_n$  be independent observations with common distribution function  $F_0$ . For any distribution function  $F$ , let  $p_i \geq 0$  be the probability associated with  $y_i$ , with  $\sum_{i=1}^n p_i = 1$ . Define  $L(F) = \prod_{i=1}^n p_i$  as the nonparametric likelihood function for  $F$ . Maximizing  $L(F)$ , subject to the constraints on  $p_i$ , yields  $p_i = n^{-1}$ . In other words, the nonparametric likelihood function attains its maximum when equal weight is placed on each observation.

Let  $\theta_0 = T(F_0)$  be a  $J$ -dimensional parameter vector for some function  $T$ . Analogous to the parametric likelihood case, inferences about  $\theta_0$  can be made using the empirical likelihood ratio  $L(F)/L(\hat{F})$ , where  $\hat{F}$  is the empirical distribution function.

Next, suppose we have  $r$  estimating functions  $g(Y; \theta) = (g_1(Y; \theta), \dots, g_r(Y; \theta))$  such that  $E_F(g(Y; \theta)) = 0$ . The main purpose of such functions is to identify the parameters of the problem. The profile empirical likelihood ratio function can then be written as

$$\mathcal{R}(\theta) = \max \left\{ \prod_{i=1}^n n p_i \mid p_i \geq 0, \sum_{i=1}^n p_i = 1, \sum_{i=1}^n p_i g(y_i; \theta) = 0 \right\}.$$

Under mild regularity conditions, it can be shown that  $-2 \log \mathcal{R}(\theta_0) \xrightarrow{d} \chi_{(J)}^2$ .<sup>2</sup> Details on the computation of the profile likelihood ratio function can be found in Owen (2001, Ch. 3.14).<sup>3</sup>

Our main focus in this paper is to compare inequality between two distinct populations.<sup>4</sup> To distinguish between the two populations, let superscripts A and B hereby indicate association with population A and B, respectively. If we let  $D_0 = (D_{1,0}, \dots, D_{J,0}) = (I_{1,0}^B - I_{1,0}^A, \dots, I_{J,0}^B - I_{J,0}^A)$ , we can test the null hypothesis that  $D_0 = D$ . Usually, applied researchers would be most interested in testing the null hypothesis that  $I_0^A = I_0^B$ , which is equivalent to testing  $D_0 = 0$ . To apply the empirical likelihood-based inference method to the generalized entropy class of inequality measures, we need to encode the parameters of our problem into suitable estimating functions.

Given that we are interested in comparing two populations, the profile empirical likelihood ratio function

<sup>2</sup>A bootstrap calibration is also possible (see Owen, 2001, Ch. 3.3)

<sup>3</sup>Computational routines for several statistical packages are available on Owen's website: <http://www-stat.stanford.edu/~owen/empirical/>

<sup>4</sup>There have been numerous studies done on empirical likelihood for the two population case (see, e.g., Wu and Yan, 2012).

is

$$\mathcal{R}(\theta^A, \theta^B) = \max \left\{ \prod_{i=1}^{n^A} n^A p_i^A \prod_{i=1}^{n^B} n^B p_i^B \left| \begin{array}{l} p_i^A \geq 0, p_i^B \geq 0, \sum_{i=1}^{n^A} p_i^A = 1, \sum_{i=1}^{n^B} p_i^B = 1, \\ \sum_{i=1}^{n^A} p_i^A g(y_i^A; \theta^A) = 0, \sum_{i=1}^{n^B} p_i^B g(y_i^B; \theta^B) = 0 \end{array} \right. \right\},$$

where  $\theta^A = (\mu^A, I^A)$ ,  $\theta^B = (\mu^B, I^A, D)$ , and the estimating functions are

$$g(Y^A; \theta^A) = \begin{pmatrix} Y_1^A - \mu_1^A \\ \vdots \\ Y_J^A - \mu_J^A \\ h_1(Y_1^A, \mu_1^A, \alpha_1) - I_1^A \\ \vdots \\ h_J(Y_J^A, \mu_J^A, \alpha_J) - I_J^A \end{pmatrix},$$

and

$$g(Y^B; \theta^B) = \begin{pmatrix} Y_1^B - \mu_1^B \\ \vdots \\ Y_J^B - \mu_J^B \\ h_1(Y_1^B, \mu_1^B, \alpha_1) - I_1^A - D_1 \\ \vdots \\ h_J(Y_J^B, \mu_J^B, \alpha_J) - I_J^A - D_J \end{pmatrix}.$$

Since we are only interested in conducting hypotheses on  $D$ , the remaining parameters in the  $\mu^A, \mu^B$  and  $I^A$  vectors are regarded as “nuisance” parameters. Following Owen (1990), we can “profile out” such parameters by maximizing over them. So the empirical likelihood ratio function for  $D$  is

$$\mathcal{R}(D) = \max_{\mu^A, \mu^B, I^A} \mathcal{R}(\mu^A, \mu^B, I^A, D).$$

To compute  $\mathcal{R}(D)$  for any vector  $D$ , we can follow Owen (1990) and use a nested algorithm which involves an “inner” and “outer” stage. The former involves solving  $\mathcal{R}(\mu^A, \mu^B, I^A, D)$  given candidate values for  $\mu^A, \mu^B$  and  $I^A$  while the latter involves maximizing  $\mathcal{R}(\mu^A, \mu^B, I^A, D)$  over the choices of  $\mu^A, \mu^B$  and  $I^A$ . The initial candidate values for such parameters could simply be the sample estimates. By maximizing over these parameter vectors, we reduce the empirical likelihood ratio function into a function of only  $D$  which is  $J$ -dimensional. Therefore, we have  $-2 \log \mathcal{R}(D_0) \xrightarrow{d} \chi_{(J)}^2$ .

## 4 Simulation evidence

To assess the size and power of our empirical likelihood-best test, we now present the results of some Monte Carlo experiments. All our hypothesis tests concern a single random variable (i.e.,  $Y_j^A = Y_k^A$  and  $Y_j^B = Y_k^B$  for all  $j, k$ ), and a single scalar measure of inequality with  $\alpha = 0$ .

To calibrate the size of our proposed method, we consider two parametric distributions: the gamma

distribution and the Singh-Maddala distribution. The cumulative distribution function for the gamma distribution is given by  $F(y) = \gamma(a_2, y/a_1)/\Gamma(a_2)$ , where  $a_1$  is a scale parameter,  $a_2$  is a shape parameter,  $\gamma(\cdot)$  is the gamma function, and  $\Gamma(\cdot)$  is the incomplete gamma function. The cumulative distribution function for the Singh-Maddala distribution is  $F(y) = 1 - (1 + b_1 y^{b_2})^{-b_3}$ , where  $b_1$  is a scale parameter, and  $b_2$  and  $b_3$  are shape parameters. Following McDonald (1984), we set  $a_2 = 2.1557$  for the gamma distribution, and  $b_2 = 1.697$  and  $b_3 = 8.368$  for the Singh-Maddala distribution which closely mimic the 1980 U.S. income distribution. The scale parameters for both distributions are set to unity.<sup>5</sup> Given these specifications, the true inequality measures under the gamma and Singh-Maddala distributions are 0.2495227 and 0.2488523, respectively.

Samples for both populations are generated from the same distribution. We test the null hypothesis that  $D_0 = 0$ , which is true. The nominal size of the test is set to 5%. Rejection frequencies for 100,000 independent trials, for sample sizes varying from  $(n^A, n^B) = (100, 100)$  to  $(n^A, n^B) = (500, 500)$ , are reported in Table 1. As a benchmark, we also report the rejection frequencies for the delta method.<sup>6</sup> Although the empirical likelihood-based method seems to over reject in small samples, it is evident that the errors in rejection probability (difference between the simulated and nominal rejection rates) subside as the sample sizes are increased. Even when we have just 300 observations on each population, these errors are less than one-half of one percentage point for both distributions.

To assess the power of our testing methodology, we conduct two Monte Carlo experiments. For the first experiment, we consider generating both samples from the gamma distribution. The second experiment involves generating both samples from the Singh-Maddala distribution. The shape parameters, of population A, for the gamma and Singh-Maddala distribution remain set to their previous values from the size simulation. We vary our choices for the shape parameters of population B so that  $D_0$ , the difference in inequality measures between population B and population A, is pre-specified to be -0.10, -0.05, 0.05, and 0.10.

As in the size simulations, we set the nominal size to 5% and test the null hypothesis that  $D_0 = 0$  (which is false in all cases here). We consider sample sizes of  $(n^A, n^B) = (100, 100)$  and  $(n^A, n^B) = (500, 500)$ , and conduct 100,000 independent trials. To correct for size distortions, the critical value of the test statistic is set equal to the 95,000th largest value observed from the size simulation. Rejection frequencies along with our choices of shape parameters (for population B) for the first and second experiment are reported in Table 2 and Table 3, respectively. We also report the (size-corrected) rejection frequencies for the delta method. From Table 2, it is apparent that the two testing methodologies perform almost identically for the first experiment (though the empirical likelihood-based method holds a slim advantage with 100 observations on each population). But when the samples are both drawn from the Singh-Maddala distribution as in Table 3, the empirical likelihood-based method clearly outperforms the delta method when  $D_0 = 0.05, 0.10$ . In fact, with 500 observations, our method has rejection frequencies that are more than five percentage points higher than those of the delta method. Overall, we can conclude that the empirical likelihood-based approach certainly matches the performance of the delta method, and in some cases outperforms it.

---

<sup>5</sup>Zheng (2001) subsequently used the Singh-Maddala distribution with the same parameter specifications to explore the asymptotic properties of decomposable poverty measure estimates with relative poverty lines. Thompson (2013) used both distributions with the same parameters to assess the size and power of his testing procedure for comparing poverty measures between two subgroups of the same population.

<sup>6</sup>The delta method can be used to obtain variances of inequality measure estimates belonging to the generalized entropy class (see, e.g., Thompson, 2010).

## 5 Empirical application

In this section, we illustrate our proposed methodology using a “real-world” application where we compare income inequality, using after-tax income data, between two populations in Canada: non-immigrants and immigrants.

We obtain our data from the Survey of Labour and Income Dynamics (SLID) for 2009. To reduce the level of heterogeneity within samples, we restrict our analysis to single (never married) individuals who reside in urban areas with a population of 500,000 or greater. For the purposes of this illustration, we consider only those immigrants who have been in Canada for 19 years or less. Our sample consists of 3,093 non-immigrants and 393 immigrants (i.e.,  $n^A = 3,093$ , and  $n^B = 393$ ).

We consider three distinct scalar measures of inequality by setting  $\alpha_1 = 0$ ,  $\alpha_2 = 0.5$ , and  $\alpha_3 = 1$ . The sample estimates of the population means and inequality measures are reported in Table 4. The nominal size of the test is set to 5%. Given three measures of inequality, the appropriate degrees of freedom for the null distribution is three. In testing  $D_0 = 0$ , our test statistic is determined to be  $-2 \log \mathcal{R}(0) = 9.851083$  which exceeds the 95th percentile of the  $\chi^2_{(3)}$  distribution (i.e., 7.814728). Thus, our null hypothesis is rejected leading us to conclude that income inequality between non-immigrants and immigrants in Canada may not be equal.

## References

- Atkinson, A. B., 1970. On the measurement of inequality. *Journal of Economic Theory* 12, 244–263.
- Biewen, M., 2002. Bootstrap inference for inequality, mobility and poverty measurement. *Journal of Econometrics* 108, 317–342.
- Cowell, F. A., 1989. Sampling variance and decomposable inequality measures. *Journal of Econometrics* 42, 27–41.
- Cowell, F. A., 2000. Measurement of inequality. In: Atkinson, A. B., Bourguignon F. (Eds.), *Handbook of Income Distribution, Volume I*. North Holland, Amsterdam.
- Cowell, F. A., 2011. *Measuring Inequality*, third ed. Oxford University Press, Oxford.
- Davidson, R., Flachaire, E., 2007. Asymptotic and bootstrap inference for inequality and poverty measures. *Journal of Econometrics* 141, 141–166.
- Kakwani, N., 1993. Statistical inference in the measurement of poverty. *The Review of Economics and Statistics* 75, 632–639.
- McDonald, J., 1984. Some generalized functions for the size distribution of income. *Econometrica* 52, 647–663.
- Owen, A. B., 1988. Empirical likelihood ratio confidence intervals for a single functional. *Biometrika* 75, 237–249.
- Owen, A. B., 1990. Empirical likelihood ratio confidence regions. *The Annals of Statistics* 18, 90–120.
- Owen, A. B., 2001. *Empirical Likelihood*. Chapman and Hall/CRC, Boca Raton.

- Thompson, B. S., 2010. Statistical inference for vector measures of inequality and poverty. *Journal of Economic Inequality* 8, 451–462.
- Thompson, B. S., 2013. Empirical likelihood-based inference for poverty measures with relative poverty lines. *Econometric Reviews* 32, 513–523.
- Wu, C., Yan, Y., 2012. Empirical likelihood inference for two-sample problems. *Statistics and its Interface* 5, 345–354.
- Zheng, B., 2001. Statistical inference for poverty measures with relative poverty lines. *Journal of Econometrics* 101, 337–356.



**Table 1**  
Rejection frequencies for size simulation

Distribution	$n^A, n^B$				
	EL				
	100,100	200,200	300,300	400,400	500,500
Gamma	5.769	5.403	5.250	5.171	5.067
Singh-Maddala	6.215	5.615	5.291	5.281	5.118
	Delta method				
Gamma	4.751	4.897	4.938	4.954	4.896
Singh-Maddala	4.705	4.982	4.919	4.986	4.908

**Table 2**  
Rejection frequencies for power simulation with gamma distributions

$D_0$	$a_2^B$	EL		$n^A, n^B$	Delta method	
		100,100	500,500		100,100	500,500
-0.10	3.501874	71.115	99.992		70.562	99.991
-0.05	2.660860	19.841	70.527		19.708	70.056
0.05	1.818233	14.964	53.615		14.865	54.100
0.10	1.576528	38.405	96.440		38.312	96.530

Notes:  $a_2^B$  is the shape parameter for population B.

**Table 3**  
Rejection frequencies for power simulation with Singh-Maddala distributions

$D_0$	$b_2^B$	$b_3^B$	EL		$n^A, n^B$	Delta method	
			100,100	500,500		100,100	500,500
-0.10	2.28	6.32	62.853	99.951		64.610	99.957
-0.05	2.01	4.45	17.243	64.067		17.913	64.098
0.05	2.09	1.38	11.143	39.256		9.118	33.982
0.10	2.08	1.17	25.120	83.727		20.146	77.966

Notes:  $b_2^B$  and  $b_3^B$  are the shape parameters for population B.

**Table 4**  
Sample estimates for empirical application

Population	$\mu$	$I_1$	$I_2$	$I_3$
Non-immigrants	25,019	0.547714	0.417437	0.378834
Immigrants	17,132	0.675557	0.501120	0.449005