Journal of Statistical
Throry & Methods
1995, Vol 6, pp 97~103

Estimation of Gini Index of the Exponential Distribution

Suk-Bok Kang[1] , Jun-Ho Kang[2] , Young-Suk Cho[1]

Abstract In this paper, we propose estimators of Gini index of the exponential distribution. We also obtain the distribution and the moments of the proposed estimators. The moments of the proposed estimators are derived by special function. We compare the maximum likelihood estimator (MLE) of Gini index with the proposed estimator of Gini index in the sense of MSE through Monte Carlo Method.

Keyword : Gini Index, Lorenz curve, MLE.

1. Introduction

The Lorenz curve proved to be a powerful tool for the analysis of a variety of scientific problems; e.g., to measure the income inequality within a population of income receivers, as a criterion to perform a partial ordering of social welfare states, to extend the concept of the Lorenz cure to functions of income. The Gini index is twice the area between the Lorenz cure and the line of equal angle. Iyengar(1960) showed that the maximum likelihood estimator (MLE) of Gini index of a lognormal distribution is asymptotically normal. Ullah and Tewari(1972) and Raghavachari(1974) obtained asymptotic expansions for the mean and variance of the MLE of Gini index of a lognormal distribution. Moothathu(1985a) derived the exact and asymptotic distributions of MLE's of Lorenz curve and Gini index of an exponential distribution. Moothathu(1985b) also derived the MLE's of Lorenz curve and Gini index of a Pareto distribution, their exact and asymptotic distributions and moments. Moothathu(1990) obtained the uniformly minimum variance unbiased estimator (UMVUE) and a strongly

---

[1] Department of Stastics, Yeungnam University, Kyungbuk, Republic of Korea
[2] College of Engineering, Kaya University, Kyungbuk, Republic of Korea

consistent asymptotically normal unbiased estimator (SCANUE) of Lorenz curve, Gini index and Theil entropy index of a Pareto distribution. We are considering the Lorenz curve and Gini index based on a single random sample from the exponential distribution with cumulative distribution function

$$F(x) = 1 - \exp\{-(x-\theta)/\sigma\},\ 0 < \theta < x,\ 0 < \sigma$$

The Lorenz curve L(p) and Gini index g of exponential distribution are given by

$$L(p) = \mu^{-1}\int_0^p F^{-1} dt$$
$$= p + \sigma(\theta+\sigma)^{-1}(1-p)\log(1-p),\ 0 \le p \le 1$$

and

$$g = 1 - 2\int_0^1 L(p)dp = \frac{\sigma}{2}(\theta+\sigma)^{-1},$$

where μ is the mean and F^{-1} is the inverse function of the cumulative distribution function.

Let $X_{(r)}(r = 1,\ 2, \cdots,\ n)$ be the r-th order statistics based on a random sample of size n $X_1, X_2, \cdots, X_n$ from F(x) and let $S = \sum_{r=1}^{n}(X_r - X_{(1)})/n$. It is well known that the joint MLE and the unrestricted best affine estimator(BAE) of (θ,σ) is give by $(X_{(1)}, S)$ and $(X_{(1)} - S/n,\ S)$, respectively. By this result, Moothathu(1985a) obtained the MLE of g is $\hat{g}_1 = S(X_{(1)} + S)^{-1}/2$.

We can propose an estimator of g by using the BAE. That is, we can propose an estimator as $\hat{g}_2 = nS/2((n-1)S + nX_{(1)})$.

The estimators of Gini index are linear function of $\hat{\lambda}_1 = S/(X_{(1)} + S)$ and $\hat{\lambda}_2 = nS/((n-1)S + nX_{(1)})$, respectively. We will find the distribution and the moment of $\hat{\lambda}_2$ in section 2. In section 3, we compare $\hat{\lambda}_1$ with $\hat{\lambda}_2$ in the sense of MSE by Monte Carlo Method.

2. Distribution and Moment of $\hat{\lambda}_i$

Sukhatme(1937) showed that $Y_{n-1} = nS/\sigma$ has a gamma distribution $G(n-1,1)$ with density function $x^{n-1}e^{-x}/\Gamma(n-1)$ and is independent of $Y_1 = n(X_{(1)} - \theta)/\sigma$, which has a gamma distribution $G(1,1)$.

Moothathu(1985a) showed that the distribution and the moment of $\hat{\lambda}_1 = S(X_{(1)} + S)^{-1}$ are given by

$$f_{\hat{\lambda}_1}(t)=\frac{A^n(1-t)^{-n}}{\Gamma(n-1)}t^{n-2}e^{-At(1-t)^{-1}}U(1,n+1,A(1-t)^{-1}),\qquad 0<t<1$$

and $$E(\hat{\lambda}_1{}^r)=\frac{1}{\Gamma(n-1)}\int_A^\infty w^{n-1}e^{-(w-A)}\int_0^{1-Aw^{-1}}t^{n+r-2}dtdw$$

$$=\frac{G(n+r-1)}{G(n-1)}\left[\frac{n}{l}(1-l)\right]^n U(n+r,n+1,n(1-l)/l)$$

where U(a,b,z) is the confluent hypergeometric function as follows

$$U(a,b,z)=\frac{1}{\Gamma(a)}\int_0^\infty t^{a-1}(1+t)^{b-a-1}e^{-zt}dt \text{ and } A=n\theta/\sigma.$$

Hence, $\hat{\lambda}_2=nS/((n-1)S+nX_{(1)})$ is distributed as $\dfrac{Y_{n-1}}{\frac{n-1}{n}Y_{n-1}+Y_1+A}$ and the joint density of Y_1 and Y_{n-1} is $f(x,y)=\dfrac{1}{\Gamma(n-1)}y^{n-1}e^{-(x+y)},\ 0<x,\ 0<y.$

From the transformation $w=A+x+\dfrac{n-1}{n}y$ and $t=\dfrac{y}{A+x+\frac{n-1}{n}y}$, the joint density of W and T is

$$f(w,t)=\frac{1}{\Gamma(n-1)}t^{n-2}w^{n-1}e^{-\left(w+\frac{wt}{n}-A\right)},\ 0<\text{t}<\frac{n}{n-1},\ A(1-\frac{n-1}{n}t)^{-1}<w.$$

Thus, we obtain the distribution of $\hat{\lambda}_2$ as follows

$$f_{\hat{\lambda}_2}(t)=e^A t^{n-2}\int_{A(1-\frac{n-1}{n}t)^{-1}}^\infty\frac{1}{\Gamma(n-1)}w^{n-1}e^{-(w+\frac{wt}{n})}dw,\quad 0<t<\frac{n}{n-1}.$$

From the result of Gradshteyn and Ryzhik(1965), the moment of $\hat{\lambda}_2$ is given by

$$E[\hat{\lambda}_2{}^r]=\frac{e^A}{\Gamma(n-1)}\int_0^{\frac{n}{n-1}}\int_{A(1-\frac{n-1}{n}t)^{-1}}^\infty w^{n-1}t^{n+r-2}e^{-(w+\frac{wt}{n})}dwdt$$

$$=\frac{e^A}{\Gamma(n-1)}\int_A^\infty\int_0^{\frac{n}{n-1}(1-\frac{A}{w})}w^{n-1}t^{n+r-2}e^{-(1+\frac{t}{n})w}dtdw$$

$$= \frac{e^A}{\Gamma(n-1)}\left[n^{n+r-1}(n+r-2)!\int_A^\infty \frac{1}{w^r}e^{-w}dw\right.$$

$$\left. -\sum_{k=0}^{n+r-2}\sum_{i=0}^{k}\frac{(n+r-2)!\,n^{n+r-1}(-A)^{k-i}}{(n-1)^k i!(k-i)!}e^{\frac{A}{n-1}}\int_A^\infty w^{-(r-i)}e^{-\frac{n}{n-1}w}dw\right]$$

$$= \frac{e^A n^{n+r-1}(n+r-2)!F_r(1)}{\Gamma(n-1)}$$

$$-\sum_{k=0}^{n+r-2}\sum_{i=0}^{k}\frac{(n+r-2)!\,n^{n+r-1}(-A)^{k-i}e^{\frac{An}{n-1}}}{\Gamma(n-1)(n-1)^k i!(k-i)!}F_{r-i}\left(\frac{n}{n-1}\right)$$

where

$$F_n(p) = \begin{cases} e^{-Ap}\sum_{k=0}^{-n}\frac{(-n)!\,A^k}{k!\,p^{-n-k+1}}, & \text{if } n<0 \\ \frac{1}{p}e^{-pA}, & \text{if } n=0 \\ -\mathrm{E}_i(-pA), & \text{if } n=1 \\ (-1)^n\frac{p^{n-1}E_i(-pA)}{(n-1)!} + \frac{e^{-pA}}{A^{n-1}}\sum_{k=0}^{n-2}\frac{(-pA)^k}{(n-1)(n-2)\cdots(n-1-k)}, & \text{if } n>1 \end{cases}$$

and $E_i(x) = C + \ln(-x) + \sum_{k=1}^{\infty}\frac{x^k}{k\cdot k!}(x<0)$, C is Euler's constant.

3. The Simulated Results

We simulate the MSE's of $\hat{g}_i$ for sample size n=5(10)95(based on 10000 Monte

Carlo runs) when the location parameter θ=0.5(1.0)2.5 and the scale parameterσ=0.5(0.5)2.0. From the following tables, the proposed estimator is more efficient than the MLE in the sense of MSE when the scale parameterσ is getting larger or the location parameter θ is getting smaller.

Table 1. The simulated MSE's $\hat{g}_i$of in the exponential distribution

	$\sigma = 0.5$					
	$\theta = 0.5$		$\theta = 1.5$		$\theta = 2.5$	
n	$MSE(\hat{g}_1)$	$MSE(\hat{g}_2)$	$MSE(\hat{g}_1)$	$MSE(\hat{g}_2)$	$MSE(\hat{g}_1)$	$MSE(\hat{g}_2)$
5	.00724	.00680	.00223	.00226	.00111	.00114
15	.00154	.00144	.00064	.00064	.00034	.00034
25	.00084	.00080	.00038	.00038	.00020	.00020
35	.00054	.00052	.00026	.00026	.00014	.00014
45	.00040	.00039	.00021	.00021	.00011	.00011
55	.00032	.00031	.00016	.00016	.00009	.00009
65	.00027	.00027	.00014	.00014	.00007	.00008
75	.00024	.00023	.00012	.00012	.00007	.00007
85	.00020	.00020	.00010	.00010	.00006	.00006
95	.00018	.00017	.00010	.00010	.00005	.00005

Table 1. (continued)

	$\sigma = 1.0$					
	$\theta = 0.5$		$\theta = 1.5$		$\theta = 2.5$	
n	$MSE(\hat{g}_1)$	$MSE(\hat{g}_2)$	$MSE(\hat{g}_1)$	$MSE(\hat{g}_2)$	$MSE(\hat{g}_1)$	$MSE(\hat{g}_2)$
5	.01072	.00911	.00500	.00491	.00283	.00286
15	.00189	.00157	.00124	.00121	.00080	.00080
25	.00093	.00079	.00069	.00068	.00046	.00046
35	.00056	.00049	.00047	.00046	.00031	.00031
45	.00041	.00036	.00036	.00035	.00024	.00024
55	.00031	.00028	.00028	.00028	.00020	.00020
65	.00025	.00023	.00024	.00024	.00016	.00016
75	.00021	.00019	.00021	.00021	.00014	.00014
85	.00018	.00016	.00018	.00018	.00012	.00012
95	.00016	.00015	.00016	.00016	.00011	.00011

Table 1. (continued)

	$\sigma = 1.5$					
	$\theta = 0.5$		$\theta = 1.5$		$\theta = 2.5$	
n	$MSE(\hat{g}_1)$	$MSE(\hat{g}_2)$	$MSE(\hat{g}_1)$	$MSE(\hat{g}_2)$	$MSE(\hat{g}_1)$	$MSE(\hat{g}_2)$
5	.01246	.01010	.00698	.00656	.00458	.00452
15	.00200	.00151	.00156	.00146	.00116	.00114
25	.00086	.00066	.00082	.00078	.00064	.00063
35	.00052	.00042	.00056	.00054	.00044	.00043
45	.00035	.00029	.00041	.00039	.00033	.00032
55	.00026	.00022	.00033	.00032	.00026	.00026
65	.00022	.00019	.00027	.00026	.00023	.00023
75	.00017	.00015	.00023	.00023	.00019	.00019
85	.00015	.00013	.00021	.00020	.00017	.00017
95	.00013	.00011	.00018	.00017	.00015	.00015

Table 1. (continued)

	$\sigma = 2.0$					
	$\theta = 0.5$		$\theta = 1.5$		$\theta = 2.5$	
n	$MSE(\hat{g}_1)$	$MSE(\hat{g}_2)$	$MSE(\hat{g}_1)$	$MSE(\hat{g}_2)$	$MSE(\hat{g}_1)$	$MSE(\hat{g}_2)$
5	.01373	.01064	.00867	.00783	.00599	.00579
15	.00205	.00144	.00175	.00158	.00138	.00132
25	.00083	.00059	.00088	.00081	.00076	.00074
35	.00048	.00036	.00056	.00052	.00051	.00050
45	.00032	.00024	.00043	.00041	.00038	.00038
55	.00024	.00019	.00033	.00031	.00031	.00031
65	.00019	.00015	.00027	.00026	.00025	.00025
75	.00015	.00012	.00023	.00022	.00022	.00022
85	.00013	.00010	.00020	.00019	.00020	.00019
95	.00011	.00009	.00018	.00017	.00017	.00017

REFERENCES

Gradsshteyn. I. S and Ryzhik. I. M(1965). *Table of Integrals, Series, and Products*, Academic Press.

Iyengar, N. S.(1960). *On the Standard Error of the Lorenz Concentration Ratio, Sankhyā*, 22, 371-378.

Moothathu, T. S. K.(1985a). *Distribution of Maximum Likelihood Estimators of*

*Lorenz Curve and Gini Index of Exponential Distribution. Ann. Inst. Statist. Math.,*37, 437- 497.

Moothathu, T. S. K.(1985b). *Sampling Distribution of Lorenz Curve and Gini Index of the Pareto Distribution. Sankhyā*, series B 47, 247-278.

Moothathu, T. S. K.(1990). *The Best Estimator of Lorenz Curve, Gini Index and Theil Entropy Index of Pareto Distribution. Sankhyā*, series B 52 ,115-127.

Raghavachari, M.(1974). A Remark on the Paper `*On Moments of Concentration Ratio* '. *Sankhyā*, series B 36, 466-467.

Sukhatme, P. V.(1937). *Test of Significance for Samples of the* χ^2 *Population with Two Degrees of Freedom*, Ann, Eugen., 8. 52-56.

Ullah, A and Tewari, V, K.(1972). *On Moments of Concentraction,ratio. Sankhyā*, series B 34, 405-408.