Inequality in Societies, Academic Institutions and Science Journals: Gini and *k*-indices

A. Ghosh, N. Chattopadhyay and B. K. Chakrabarti

Saha Institute of Nuclear Physics Econophys-Kolkata VIII

March 15, 2014

arXiv:1401.6951 (2014)

Social inequality refers to relational processes in society that have the effect of limiting or harming a groups social status, social class, and social circle (wikipedia).

Inequality measures:

- Parade of Dwarfs (Range R, Relative Mean Deviation *M*)
- Frequency distribution (Variance V, Coefficient of variation c)
- Lorenz Curve (Gini Coefficient g)
- Log transformation (Log variance v)

Our interest is to measure inequality of citation for institutions and journals using Lorenz Curve (Gini Coefficient g; *k*-indices)



w = Cumulative fraction citations (income) of any institutions/journals (country).

n = Cumulative fraction papers (people) of the institutions/journals (country) from lowest to highest citation (income).

gini index
$$g = \frac{A}{(A+B)}$$

k-index value gives that 1 - k fraction of papers possess more citations than k fraction of papers.



Data collection: SINP-1990

WEB OF SCIENCE™				i neumens"	
Search Web of Science TM Core Collection		My Tools 🔻	Search History	Marked List	
	We	lcome to the new	Web of Science! View	a brief tutorial.	
Basic Search Saha Inst Nuclear Physics View Abbreviations List + Add Another Field	Address	Search	Click her improve	re for tips to your search.	



Data collection: SINP-1990



Citation distribution: SINP-1990

papers	citation	papers	citation
12	0	1	11
18	1	1	12
6	2	2	14
6	3	1	15
11	4	1	17
4	5	1	18
3	6	2	34
5	7	1	38
5	8	1	50
3	9	1	61
5	10	1	66

Result: SINP-1990



cumulative fraction papers from lowest to highest citations

 $g = \frac{A}{A+B} \simeq 0.72$

Results

Institution/University	Year	total	g	k
		papers/citations		
	1980	866/16107	0.67	0.75
University of Melbourne	1990	1131/30349	0.68	0.75
Oniversity of Melbourne	2000	2116/57871	0.65	0.74
	2010	5255/63151	0.68	0.75
University of Tokyo	1980	2871/60682	0.69	0.76
	1990	4196/108127	0.68	0.76
	2000	7955/221323	0.70	0.76
	2010	9154/91349	0.70	0.76
Harvard University	1980	4897/225626	0.73	0.78
	1990	6036/387244	0.73	0.78
	2000	9566/571666	0.71	0.77
	2010	15079/263600	0.69	0.76

Table: *g*-index and *k*-index values for papers and citations (up to December 2013).

Institution/University	Year	total	g	k
		papers/citations		
	1980	32/170	0.72	0.74
Saba Instituto	1990	91/666	0.66	0.73
of Nuclear Division	2000	148/2225	0.77	0.79
of Nuclear Physics	2010	238/1896	0.71	0.76
Indian Institute of Science	1980	450/4728	0.73	0.78
	1990	573/8410	0.70	0.76
	2000	874/19167	0.67	0.75
	2010	1624/11497	0.62	0.73
	1980	167/2024	0.70	0.76
Tata Institute of Fundamental Research	1990	303/4961	0.73	0.77
	2000	439/11275	0.74	0.77
	2010	573/9988	0.78	0.79

Results

Journals	Year	total	g	k
		papers/citations		
	1980	2904/178927	0.80	0.81
Naturo	1990	3676/307545	0.86	0.85
Nature	2000	3021/393521	0.81	0.82
	2010	2577/100808	0.79	0.81
Science	1980	1722/111737	0.77	0.80
	1990	2449/228121	0.84	0.84
	2000	2590/301093	0.81	0.82
	2010	2439/85879	0.76	0.79
	1980	-	-	-
PNAS	1990	2133/282930	0.54	0.70
	2000	2698/315684	0.49	0.68
	2010	4218/116037	0.46	0.66

Table: *g*-index and *k*-index values for papers and citations (up to December 2013).

Journals	Year	total	g	k
		papers/citations		
	1980	394/72676	0.54	0.70
Call	1990	516/169868	0.50	0.68
Cell	2000	351/110602	0.56	0.70
	2010	573/32485	0.68	0.75
Physical Review Letter	1980	1196/87773	0.66	0.74
	1990	1904/156722	0.63	0.74
	2000	3124/225591	0.59	0.72
	2010	3350/73917	0.51	0.68
Physical Review A	1980	639/24802	0.61	0.73
	1990	1922/54511	0.61	0.72
	2000	1410/38948	0.60	0.72
	2010	2934/26314	0.53	0.69

Country	g	k	Country	g	k
Brazil	0.62	0.73	Columbia	0.55	0.70
Denmark	0.36	0.63	Finland	0.47	0.67
India	0.45	0.66	Indonesia	0.44	0.65
Japan	0.31	0.61	Kenya	0.61	0.73
Malaysia	0.50	0.68	Netherlands	0.44	0.66
NewZeland	0.37	0.63	Norway	0.36	0.63
Panama	0.44	0.66	SriLanka	0.40	0.65
Sweden	0.38	0.64	Tanzania	0.53	0.70
Tunisia	0.50	0.69	United Kingdom	0.36	0.63
Uruguay	0.49	0.68	-	-	-

Table: *g*-index and *k*-index values for income distribution of various countries of the world during the years 19963 to 1983 as obtained analyzing data reported in A. F. Shorrocks, *Ranking income distributions*, Economica **50**, pp. 3-17 (1983).



Figure: Plot of 1 - w against 1 - n for citation distributions for a few institutions, showing that their variation with the corresponding publication numbers follow a Pareto type power law behavior beyond the *k*-index value of *n*: $1 - w \sim (1 - n)^{\alpha}$ for $n \geq k$, with $\alpha = 0.50 \pm 0.10$.



Figure: Plot of 1 - w against 1 - n for citation distributions for a few science journals, showing that their variation with the corresponding publication numbers follow a Pareto type power law behavior beyond the *k*-index value of *n*: $1 - w \sim (1 - n)^{\alpha}$ for $n \geq k$, with $\alpha = 0.50 \pm 0.10$.

- Strong inequality and universality ($g = 0.70 \pm 0.07$) across all the universities and institutions of the world, while for the journals we find $g = 0.65 \pm 0.15$ for any typical year.
- *k*-index value for citation has a value around 0.75 ± 0.05 for different universities and institutions across the world and around 0.77 ± 0.10 for the science journals.
- We define a new inequality measure, k-index, saying that the cumulative citations of (1 k) fraction of papers exceed those earned by the fraction (k) of the publications.

Thank You