Temporal changes in the parameters of statistical distribution of journal impact factor

Mishra, SK
North-Eastern Hill University, Shillong (India)

10. March 2010

Temporal Changes in the Parameters of Statistical Distribution of Journal Impact Factor

SK Mishra
Dept. of Economics
North-Eastern Hill University
Shillong, Meghalaya (India)
Contact: mishrasknehu@yahoo.com

1. Introduction: The Journal Impact Factor (JIF) is one of the very important numerical measures of scientific or research importance of a journal. The importance or quality of a paper/article (and, by implication, the author(s) of the paper/article) published in a journal is often judged by the JIF of the journal concerned. Impact factors are calculated every year for those journals that are indexed in Thomson Reuter's Journal Citation Reports.

Statistical distribution of Journal Impact Factor (JIF) is characteristically asymmetric and non-mesokurtic. Even the distribution of $\log_{10}$(JIF) exhibits conspicuous skewness and non-mesokurticity, characterizing Pearson's type-IV distribution (Mishra, 2009). In view of this observation, statistical distributions such as Burr, Dagum, Johnson S_U, log-logistic, etc. fit quite well to the $\log_{10}$(JIF) data (Mishra, 2010_a). However, it has been found that although Burr-XII, Dagum and Johnson S_U distributions fit better to the $\log_{10}$(JIF) than any other distribution, the estimated parameters of Burr-XII and Dagum distributions do not exhibit stability over the samples. On the other hand, the estimated parameters of Johnson S_U exhibit stability over the samples (Mishra, 2010_b). In view of this, Johnson S_U appears to be the best choice to fit to the $\log_{10}$(JIF) data.

The Johnson system is based on the principle of translation of a given statistical distribution such that the resulting (post-translation) distribution is a normal distribution (Johnson, 1949; Tadikamalla, 1980). It provides a unique distribution corresponding to each pair of mathematically possible values of skewness and kurtosis (George, 2007). It comprises three families of distributions: Johnson S_U, Johnson S_B and Johnson S_L. The probability density function (pdf) of Johnson S_U distribution is given as:

$$f(x) = \frac{\delta}{\lambda\sqrt{2\pi}\sqrt{z^2+1}} \exp\left(-\frac{1}{2}\left(\gamma + \delta \ln\left(z + \sqrt{z^2+1}\right)\right)^2\right), \text{where}$$

γ is the shape parameter

ζ is the location parameter

$\delta > 0$ is another shape parameter

$z = (x - \zeta)/\lambda; \quad x: -\infty < x < +\infty,$

$\lambda > 0$ is the scale parmeter

There are three well known methods of estimation of the parameters of Johnson S_U (as well as SL and SB) distribution: (1) the moments matching method (Draper, 1952), (2) the percentile matching method (Slifker and Shapiro, 1980), and (3) the quantile estimation method (Wheeler, 1980). Another method, namely the MLE-least squares, suggested by George (2007), also performs very well.

From the given univariate sample data $x = (x_1, x_2, ..., x_n)$, which may have skew and non-mesokurtic distribution, the parameters $(\gamma, \delta, \lambda, \varsigma)$ may be estimated by any suitable method. Using these estimated parameters one may transform the given sample data, x, into $y = (y_1, y_2, ..., y_n)$ such that

$y_i = \gamma + \delta \log_e\{z_i + sqrt(z_i^2 + 1)\} = \gamma + \delta \sinh^{-1}(z_i) : z_i = (x_i - \varsigma)/\lambda;\ i = 1, 2, ..., n.$ The resulting variate, y, is normally distributed (George, 2007).

Fig.-1: Effect of Change in Location Parameter (ζ)

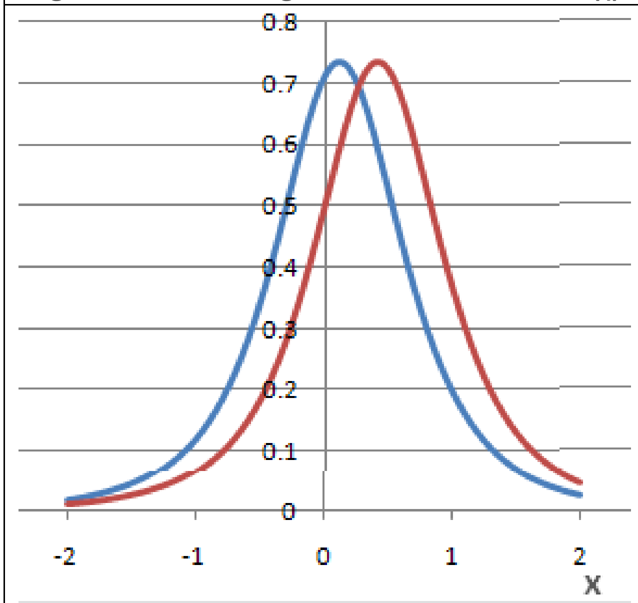


Fig.-2: Effect of Change in Scale Parameter (λ)

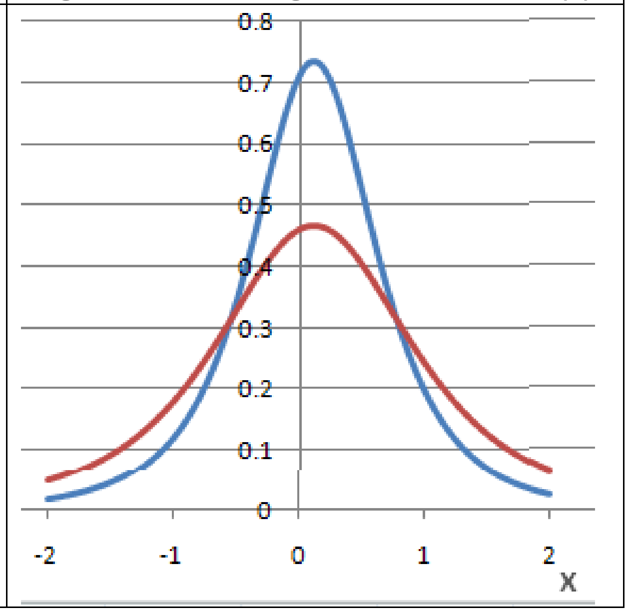


Fig.-3: Effect of Change in Shape Parameter (δ)

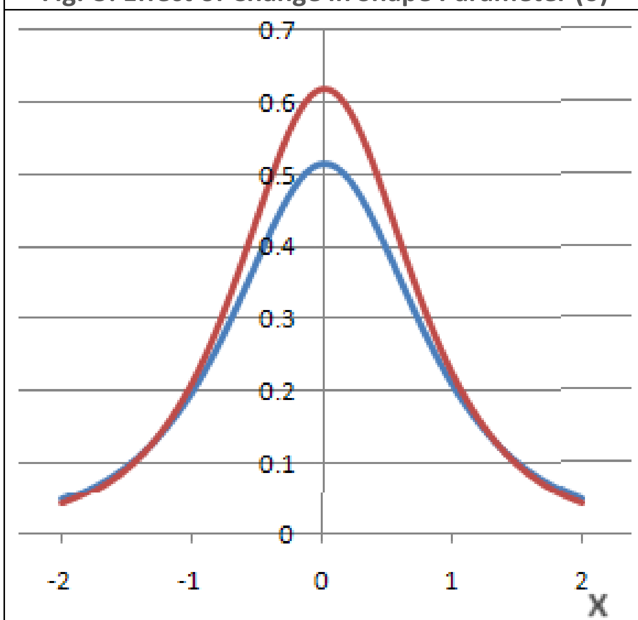
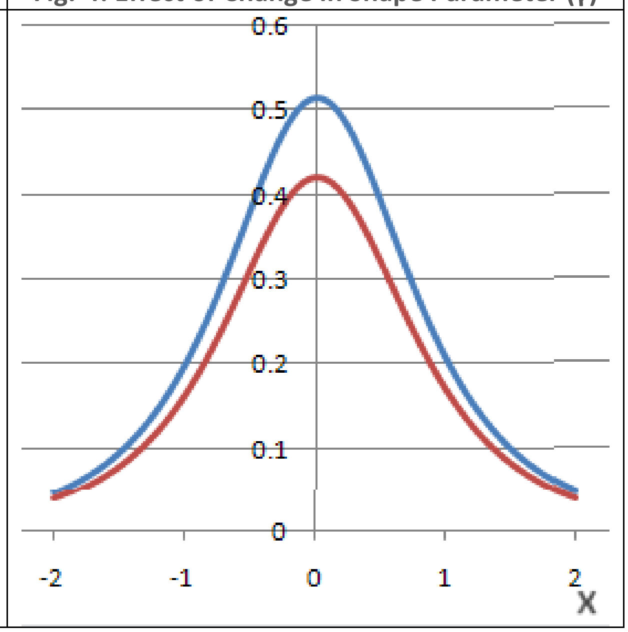


Fig.-4: Effect of Change in Shape Parameter (γ)



A visual aid to understanding the meanings of the four parameters of Johnson S_U distribution is provided in the four figures (Fig.-1 through Fig.-4). These figures depict the effect of increase in one of the parameters, keeping the other three constant. As may be seen in Fig.-1, an increase in the value of zeta (ζ), the location parameter, pushes the pdf curve to the right, indicating that the density (as well as peak) has shifted further right to zero and thus increasing the degree of negative skewness. An increase in the scale parameter (λ) makes the distribution less dense in the central region around the peak and more dispersed over longer distance in both the sides of zero (Fig.-2). Fig.-3 and Fig.-4 show the effects of an increase in the two shape parameters, Delta (δ) and gamma (γ), while zeta and lambda are fixed at zero and unity respectively**.** In matters of kurtocity, they have opposite effects. However, an increase in delta is more effective in the central region around the peak, not affecting the tails region much, though elongating the tails and thickening them slightly.

2. The Objectives and the Database: The objectives of this paper are to estimate the parameters of Johnson S_U distribution fitting to the $\log_{10}$(JIF) data for several years and study the temporal variations in those estimated parameters. We also study 'over-the-samples stability' in the estimated parameters for each year. We have used positive (non-zero) JIF data for the years 2001 (5679 journals), 2002 (5475 journals), 2003 (5702 journal), 2004 (5913 journals), 2005 (6033 journals), 2006 (6152 journals), 2007 (6226 journals) and 2008 (6545 journals). The sources of data are:

(1) For the Year 2001: www.genebee.msu.su/journals/if01a.html
(2) For the Year 2002: www.genebee.msu.su/journals/if02a.html
(3) For the Year 2003: www.genebee.msu.su/journals/if03a.html
(4) For the Year 2004: www.pmf.ukim.edu.mk/PMF/Chemistry/PDF/IF_2004.pdf
(5) For the Year 2005: http://gezhi.org/wp-content/uploads/2006/06/2005%20SCI.zip
(6) For the Year 2006: http://www.cricyt.edu.ar/secedoc/fi/fi2006.pdf
(7) For the Year 2007: http://www.icast.org.in/FACTOR.html
(8) For the Year 2008: http://www.mazums.ac.ir/files/f27302Impact%20Factor%202008.pdf

From the pdf files the JIF data were extracted by writing suitable computer programs.

3. The Methodology: Since one of our objectives in this paper is to study 'over-the-samples' variations in the estimated parameters of Johnson S_U distribution fitted to the $\log_{10}$(JIF) data, we have not used the entire set of data (for any particular year) for fitting the distribution and thus estimating its parameters. Instead, we have drawn 30 samples (with replacement), each of the size 5000, from each year's data. We have fitted the distribution to each sample (for each year). In a sense, it is a sort of re-sampling close to bootstrapping. We have not fixed any particular percentage for sampling; it is obvious that 5000 makes 88% of 5679 for 2001 JIF data, while it makes only 76.4% of 6545 for 2008 JIF data. But, in any case, we hold that the sub-sample size is large enough to represent the entire data for any year.

4. The Results: The results of our analysis are presented in tables 1.1 through 8.2. The tables 1.1, 2.1, … , 8.1 present the estimated parameters ($\gamma, \delta, \lambda, \varsigma$) for 30 samples (S_j; j=1,2, …, 30) for the years 2001, 2002, … , 2008 respectively. Tables 1.2, 2.2, … , 8.2 present the descriptive statistics (median, mean, etc) obtained from the estimated parameters reported in the tables 1.1, 1.2, …. , 8.1 respectively. It may be observed that the standard errors of estimate of mean of different parameters ($\gamma, \delta, \lambda, \varsigma$) are quite small and the spread of parameters between -95% and +95% confidence intervals is quite narrow (Fig. 6 through Fig.-9). Median values of parameters are very close to the mean values showing symmetry in variation around the mean values. All these statistics indicate over-the-samples stability in the estimated parameters and suitability of Johnson S_U distribution to the data for all the years. Although we do not intend to report the details here, we have found, nevertheless, that other distributions such as Dagum and Burr fit extremely well to the $\log_{10}$(JIF) data in all the sub-samples (for all the 8 years), but their parameters do not exhibit stability over the sub-samples.

As depicted in Fig.-5, the mean value of zeta over the years has an increasing trend. This is also borne out by the estimated values of skewness, increasing since 2003, as presented in Table 9. The mean value of lambda as well as delta is decreasing over time, indicating growing concentration in the central region around the peak, but with a moderation effected by the increasing mean value of gamma. This is in consonance with the finding that kurtosis of the $\log_{10}$(JIF) distribution is increasing since 2004, as reported in Table-9. It also reconfirms that the $\log_{10}$(JIF) distribution is Pearson's type-IV.

Table-1.1: Estimated Parameters of Johnson S_U Distribution fitted to Thirty Sub-Samples (Size=5000) of JIF data for 2001

S_j	γ	δ	λ	ζ	S_j	γ	δ	λ	ζ	S_j	γ	δ	λ	ζ
1	0.59153	2.25880	0.95188	0.18038	11	0.69241	2.35240	0.99174	0.22795	21	0.64014	2.19400	0.91960	0.19984
2	0.63347	2.25020	0.94412	0.20136	12	0.63590	2.28310	0.95563	0.20193	22	0.58056	2.17790	0.90978	0.17449
3	0.60922	2.17490	0.90288	0.18549	13	0.63734	2.20000	0.92291	0.20307	23	0.62623	2.20800	0.91879	0.19568
4	0.60732	2.13900	0.88614	0.18412	14	0.60081	2.23530	0.93325	0.18603	24	0.63699	2.27360	0.95579	0.20379
5	0.59797	2.23810	0.94840	0.18815	15	0.65072	2.24290	0.93370	0.20596	25	0.63690	2.19250	0.91157	0.19816
6	0.59823	2.17280	0.90778	0.18453	16	0.58357	2.25810	0.94664	0.17741	26	0.67103	2.25190	0.94002	0.21294
7	0.62713	2.17740	0.90423	0.19589	17	0.63288	2.25360	0.95019	0.20095	27	0.60619	2.23720	0.93819	0.18975
8	0.64695	2.31920	0.96684	0.20451	18	0.57264	2.25670	0.94298	0.17092	28	0.66167	2.23100	0.92758	0.21212
9	0.60682	2.21790	0.93172	0.18880	19	0.59870	2.21090	0.92548	0.18605	29	0.65840	2.29440	0.95885	0.20867
10	0.64576	2.25690	0.94165	0.20696	20	0.59728	2.19470	0.91617	0.18477	30	0.61366	2.21020	0.92733	0.19184

.

Table-1.2: Descriptive Statistics of Estimated Parameters of Johnson S_U Distribution fitted to Thirty Sub-Samples (Size=5000) of JIF data for 2001

Parameters	Median	Mean	Std. Error	Conf. -95%	Conf. 95%	Minimum	Maximum	Std. Dev.
γ	0.626680	0.623281	0.005247	0.612549	0.634013	0.572640	0.692410	0.028741
δ	2.236250	2.232120	0.008495	2.214745	2.249495	2.139000	2.352400	0.046531
λ	0.933475	0.933728	0.004037	0.925470	0.941985	0.886140	0.991740	0.022113
ζ	0.195785	0.195084	0.002342	0.190294	0.199874	0.170920	0.227950	0.012828

.

Table-2.1: Estimated Parameters of Johnson S_U Distribution fitted to Thirty Sub-Samples (Size=5000) of JIF data for 2002

S_j	γ	δ	λ	ζ	S_j	γ	δ	λ	ζ	S_j	γ	δ	λ	ζ
1	0.42938	2.19120	0.91359	0.12570	11	0.44399	2.30560	0.96891	0.13427	21	0.47109	2.22230	0.93078	0.14591
2	0.50511	2.23190	0.92921	0.16117	12	0.43546	2.17480	0.91239	0.13010	22	0.41185	2.21860	0.93144	0.11632
3	0.44963	2.20620	0.92520	0.13274	13	0.46925	2.24800	0.94975	0.14810	23	0.46077	2.21080	0.92583	0.14263
4	0.42167	2.21590	0.93332	0.12943	14	0.44249	2.18580	0.91803	0.13536	24	0.37619	2.20410	0.92526	0.10306
5	0.43201	2.19230	0.91735	0.13399	15	0.47527	2.18070	0.89890	0.14292	25	0.45967	2.31390	0.97585	0.14240
6	0.44690	2.27840	0.96110	0.13800	16	0.48216	2.19300	0.91636	0.15242	26	0.41005	2.18450	0.91046	0.12070
7	0.43721	2.18740	0.91323	0.13073	17	0.46135	2.18850	0.91378	0.14311	27	0.43766	2.20250	0.92359	0.13460
8	0.46610	2.27730	0.95227	0.14152	18	0.45497	2.18880	0.91937	0.13846	28	0.45529	2.20710	0.92812	0.13968
9	0.48807	2.19010	0.91461	0.14974	19	0.43003	2.21230	0.92256	0.12406	29	0.42107	2.15330	0.89767	0.12585
10	0.39887	2.14820	0.89324	0.11036	20	0.47574	2.20180	0.91555	0.14765	30	0.43616	2.19490	0.91401	0.12700

.

Table-2.2: Descriptive Statistics of Estimated Parameters of Johnson S_U Distribution fitted to Thirty Sub-Samples (Size=5000) of JIF data for 2002

Parameters	Median	Mean	Std. Error	Conf. -95%	Conf. 95%	Minimum	Maximum	Std. Dev.
γ	0.445445	0.446182	0.005094	0.435764	0.456600	0.376190	0.505110	0.027899
δ	2.202150	2.210340	0.007182	2.195651	2.225029	2.148200	2.313900	0.039339
λ	0.920965	0.925058	0.003570	0.917757	0.932359	0.893240	0.975850	0.019553
ζ	0.134980	0.134933	0.002299	0.130231	0.139634	0.103060	0.161170	0.012591

.

Table-3.1: Estimated Parameters of Johnson S_U Distribution fitted to Thirty Sub-Samples (Size=5000) of JIF data for 2003

S_j	γ	δ	λ	ζ	S_j	γ	δ	λ	ζ	S_j	γ	δ	λ	ζ
1	0.35235	2.13040	0.87264	0.12783	11	0.36410	2.21050	0.90946	0.12492	21	0.39546	2.28110	0.93839	0.14282
2	0.34576	2.22190	0.92221	0.11902	12	0.37131	2.22910	0.92179	0.13410	22	0.30950	2.17140	0.89260	0.10764
3	0.36974	2.15340	0.88959	0.13456	13	0.40932	2.17800	0.89583	0.15124	23	0.32481	2.22460	0.91880	0.11248
4	0.39417	2.20810	0.91278	0.13763	14	0.35229	2.24320	0.93071	0.12601	24	0.31832	2.17900	0.89659	0.11061
5	0.33678	2.22610	0.91425	0.11663	15	0.34878	2.11490	0.86095	0.12167	25	0.36259	2.18520	0.89995	0.12677
6	0.34739	2.19930	0.90668	0.12202	16	0.38851	2.22300	0.91812	0.13769	26	0.33432	2.21020	0.91049	0.11745
7	0.37820	2.27180	0.93864	0.13511	17	0.34202	2.19220	0.90042	0.12162	27	0.30878	2.17720	0.90136	0.10629
8	0.33218	2.19380	0.90838	0.11392	18	0.38275	2.18260	0.90756	0.14053	28	0.35745	2.20580	0.91977	0.12516
9	0.35223	2.21120	0.91748	0.12187	19	0.33304	2.12950	0.87952	0.11873	29	0.31681	2.19200	0.90135	0.10712
10	0.37202	2.23590	0.92482	0.13085	20	0.33912	2.26600	0.93045	0.11633	30	0.34174	2.18130	0.90195	0.11875

.

Table-3.2: Descriptive Statistics of Estimated Parameters of Johnson S_U Distribution fitted to Thirty Sub-Samples (Size=5000) of JIF data for 2003

Parameters	Median	Mean	Std. Error	Conf. -95%	Conf. 95%	Minimum	Maximum	Std. Dev.
γ	0.350505	0.352728	0.004771	0.342970	0.362486	0.308780	0.409320	0.026131
δ	2.202550	2.200957	0.007218	2.186195	2.215719	2.114900	2.281100	0.039533
λ	0.908920	0.908118	0.003296	0.901376	0.914859	0.860950	0.938640	0.018055
ζ	0.121945	0.124246	0.002059	0.120034	0.128457	0.106290	0.151240	0.011278

.

Table-4.1: Estimated Parameters of Johnson S_U Distribution fitted to Thirty Sub-Samples (Size=5000) of JIF data for 2004

S_j	γ	δ	λ	ζ	S_j	γ	δ	λ	ζ	S_j	γ	δ	λ	ζ
1	0.45093	2.31730	0.96501	0.18122	11	0.51563	2.33470	0.98488	0.21089	21	0.41489	2.40910	1.01580	0.16633
2	0.45800	2.28890	0.96135	0.19135	12	0.45899	2.38910	1.00060	0.18336	22	0.45759	2.42840	1.01990	0.18725
3	0.45482	2.27210	0.95620	0.18771	13	0.46537	2.41290	1.02730	0.19507	23	0.42719	2.33690	0.98512	0.16878
4	0.46433	2.28490	0.95115	0.18890	14	0.40657	2.39580	1.01670	0.16199	24	0.42445	2.36160	0.99498	0.17372
5	0.42898	2.41570	1.01490	0.17855	15	0.47238	2.36070	0.98856	0.19217	25	0.41685	2.33320	0.98434	0.17228
6	0.44256	2.27530	0.95067	0.17666	16	0.42414	2.37110	1.00780	0.17557	26	0.42899	2.27990	0.95587	0.17900
7	0.47852	2.36240	0.99502	0.19923	17	0.40068	2.30900	0.96536	0.16209	27	0.45399	2.35820	0.99377	0.18486
8	0.43916	2.27450	0.95190	0.17706	18	0.42430	2.24030	0.93294	0.17251	28	0.41820	2.32350	0.98405	0.17354
9	0.47026	2.37500	1.00260	0.19109	19	0.44546	2.38320	1.00540	0.18806	29	0.43346	2.25650	0.93560	0.17267
10	0.43492	2.29080	0.95996	0.17331	20	0.42764	2.26340	0.94601	0.17397	30	0.45264	2.30780	0.97211	0.18749

.

Table-4.2: Descriptive Statistics of Estimated Parameters of Johnson S_U Distribution fitted to Thirty Sub-Samples (Size=5000) of JIF data for 2004

Parameters	Median	Mean	Std. Error	Conf. -95%	Conf. 95%	Minimum	Maximum	Std. Dev.
γ	0.440860	0.443063	0.004488	0.433885	0.452241	0.400680	0.515630	0.024580
δ	2.333950	2.333740	0.009837	2.313621	2.353859	2.240300	2.428400	0.053879
λ	0.984610	0.980862	0.004916	0.970807	0.990917	0.932940	1.027300	0.026928
ζ	0.178775	0.180889	0.002036	0.176724	0.185054	0.161990	0.210890	0.011154

.

.

Table-5.1: Estimated Parameters of Johnson S_U Distribution fitted to Thirty Sub-Samples (Size=5000) of JIF data for 2005														
S_j	γ	δ	λ	ζ	S_j	γ	δ	λ	ζ	S_j	γ	δ	λ	ζ
1	0.44193	2.38030	0.97902	0.20477	11	0.48545	2.49900	1.02970	0.23031	21	0.44337	2.37190	0.97325	0.20735
2	0.42969	2.26020	0.92594	0.20114	12	0.46293	2.29850	0.93155	0.21954	22	0.39562	2.37260	0.97485	0.19391
3	0.42411	2.32740	0.95614	0.20117	13	0.43713	2.33590	0.94923	0.20784	23	0.51179	2.42560	0.99284	0.23817
4	0.38390	2.39900	0.98385	0.18059	14	0.42022	2.30180	0.94449	0.20433	24	0.47622	2.30930	0.93777	0.22260
5	0.48043	2.35040	0.95662	0.22618	15	0.42481	2.31270	0.93849	0.19941	25	0.45840	2.27330	0.92537	0.21513
6	0.41935	2.24670	0.91010	0.19904	16	0.46865	2.44010	1.00460	0.22101	26	0.43473	2.22680	0.90108	0.20781
7	0.45445	2.39950	0.98642	0.21364	17	0.44269	2.29780	0.94515	0.20898	27	0.47445	2.25420	0.91730	0.22246
8	0.46689	2.33800	0.96214	0.22005	18	0.45544	2.35270	0.96050	0.21423	28	0.44351	2.41090	0.99268	0.21231
9	0.44429	2.41030	0.98878	0.20985	19	0.50626	2.36760	0.96689	0.23608	29	0.44016	2.30590	0.94358	0.20639
10	0.47034	2.26620	0.92340	0.22437	20	0.48861	2.30360	0.93881	0.22794	30	0.46274	2.36230	0.96473	0.21595

.

Table-5.2: Descriptive Statistics of Estimated Parameters of Johnson S_U Distribution fitted to Thirty Sub-Samples (Size=5000) of JIF data for 2005								
Parameters	Median	Mean	Std. Error	Conf. -95%	Conf. 95%	Minimum	Maximum	Std. Dev.
γ	0.449370	0.451619	0.005357	0.440663	0.462574	0.383900	0.511790	0.029339
δ	2.336950	2.340017	0.011711	2.316066	2.363968	2.226800	2.499000	0.064142
λ	0.956380	0.956842	0.005487	0.945620	0.968065	0.901080	1.029700	0.030054
ζ	0.212975	0.213085	0.002329	0.208322	0.217848	0.180590	0.238170	0.012756

.

Table-6.1: Estimated Parameters of Johnson S_U Distribution fitted to Thirty Sub-Samples (Size=5000) of JIF data for 2006														
S_j	γ	δ	λ	ζ	S_j	γ	δ	λ	ζ	S_j	γ	δ	λ	ζ
1	0.56413	2.37530	0.96303	0.29145	11	0.51293	2.39030	0.98885	0.26758	21	0.56224	2.42920	0.99368	0.28683
2	0.47144	2.17360	0.87449	0.24676	12	0.53629	2.20330	0.88101	0.27180	22	0.45975	2.20980	0.88514	0.23724
3	0.44962	2.18280	0.87828	0.23651	13	0.56653	2.47220	1.00120	0.28910	23	0.52638	2.17300	0.86219	0.26793
4	0.51442	2.40840	0.98237	0.27094	14	0.43406	2.15950	0.86994	0.22835	24	0.56528	2.29730	0.91311	0.28295
5	0.49031	2.15750	0.85858	0.25766	15	0.46637	2.23820	0.90715	0.24652	25	0.47523	2.43150	1.00250	0.24632
6	0.63406	2.39200	0.96153	0.31200	16	0.48152	2.18180	0.88274	0.25097	26	0.52083	2.22480	0.89392	0.26510
7	0.42737	2.21850	0.88999	0.22660	17	0.46402	2.23930	0.90110	0.24087	27	0.49360	2.21150	0.89299	0.24858
8	0.50527	2.19030	0.88469	0.26156	18	0.49017	2.25220	0.90665	0.24831	28	0.46551	2.38710	0.98327	0.24191
9	0.51451	2.18720	0.87008	0.26370	19	0.48296	2.26470	0.92440	0.25568	29	0.60251	2.40070	0.97047	0.30203
10	0.51298	2.26270	0.91088	0.26526	20	0.51354	2.12440	0.83919	0.26278	30	0.44211	2.19420	0.87983	0.23757

.

Table-6.2: Descriptive Statistics of Estimated Parameters of Johnson S_U Distribution fitted to Thirty Sub-Samples (Size=5000) of JIF data for 2006								
Parameters	Median	Mean	Std. Error	Conf. -95%	Conf. 95%	Minimum	Maximum	Std. Dev.
γ	0.499435	0.504865	0.009027	0.486402	0.523327	0.427370	0.634060	0.049444
δ	2.231500	2.267777	0.018599	2.229738	2.305816	2.124400	2.472200	0.101870
λ	0.897510	0.915108	0.008936	0.896833	0.933384	0.839190	1.002500	0.048943
ζ	0.259610	0.260362	0.003911	0.252364	0.268360	0.226600	0.312000	0.021420

.

Table-7.1: Estimated Parameters of Johnson S_U Distribution fitted to Thirty Sub-Samples (Size=5000) of JIF data for 2007														
S_j	γ	δ	λ	ζ	S_j	γ	δ	λ	ζ	S_j	γ	δ	λ	ζ
1	0.56861	2.21910	0.85870	0.29164	11	0.52034	2.12700	0.82385	0.28035	21	0.48976	2.08110	0.81275	0.26180
2	0.59287	2.24400	0.88119	0.30932	12	0.46157	2.13910	0.83388	0.25571	22	0.52804	2.13910	0.82722	0.27777
3	0.49355	2.13490	0.83490	0.26928	13	0.47956	2.20500	0.86366	0.25406	23	0.53623	2.16520	0.84867	0.28508
4	0.48582	2.16320	0.84251	0.26142	14	0.50228	2.11420	0.83000	0.27021	24	0.48054	2.15640	0.84098	0.26206
5	0.55918	2.20940	0.87029	0.29501	15	0.47077	2.15520	0.84223	0.25380	25	0.54671	2.21810	0.86938	0.28822
6	0.49858	2.13890	0.83258	0.26572	16	0.51600	2.17390	0.85346	0.27206	26	0.56770	2.18610	0.85962	0.29982
7	0.49825	2.13940	0.83464	0.26618	17	0.58819	2.13550	0.82705	0.30278	27	0.48640	2.15000	0.84386	0.26005
8	0.54031	2.22920	0.87337	0.28318	18	0.51535	2.22860	0.87435	0.27537	28	0.50569	2.22930	0.87816	0.26775
9	0.49397	2.14910	0.83774	0.26383	19	0.48989	2.18910	0.85184	0.26717	29	0.55364	2.14820	0.83124	0.29273
10	0.50158	2.23900	0.88488	0.26974	20	0.51480	2.11370	0.81987	0.27586	30	0.51002	2.22640	0.88090	0.27631

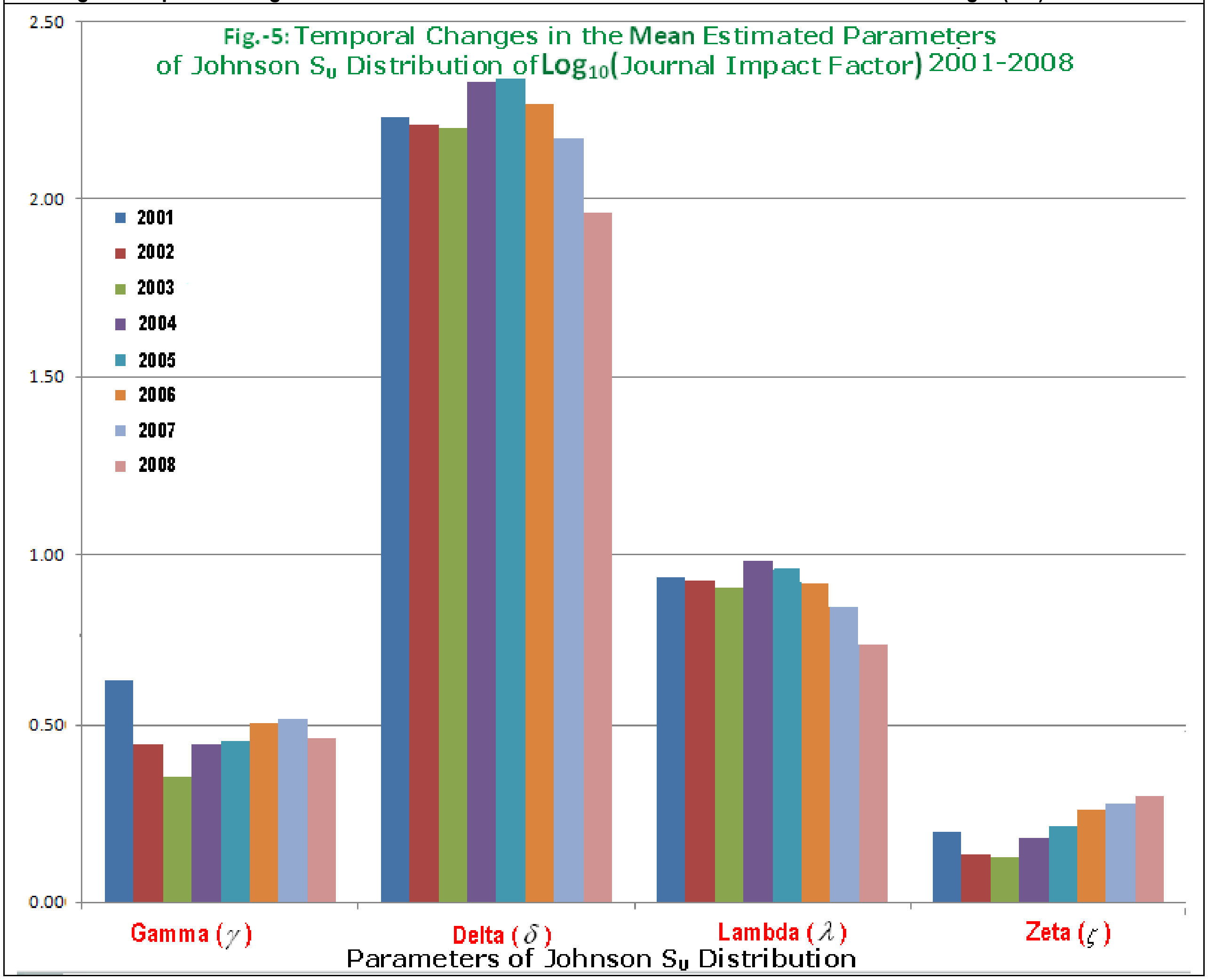
Table-7.2: Descriptive Statistics of Estimated Parameters of Johnson S_U Distribution fitted to Thirty Sub-Samples (Size=5000) of JIF data for 2007								
Parameters	Median	Mean	Std. Error	Conf. -95%	Conf. 95%	Minimum	Maximum	Std. Dev.
γ	0.507855	0.516540	0.006328	0.503597	0.529483	0.461570	0.592870	0.034661
δ	2.159800	2.171580	0.008023	2.155172	2.187988	2.081100	2.244000	0.043942
λ	0.843185	0.848792	0.003773	0.841077	0.856508	0.812750	0.884880	0.020663
ζ	0.271135	0.275143	0.002721	0.269577	0.280708	0.253800	0.309320	0.014905

Table-8.1: Estimated Parameters of Johnson S_U Distribution fitted to Thirty Sub-Samples (Size=5000) of JIF data for 2008														
S_j	γ	δ	λ	ζ	S_j	γ	δ	λ	ζ	S_j	γ	δ	λ	ζ
1	0.43882	2.02260	0.74972	0.28192	11	0.47423	1.92100	0.70942	0.29591	21	0.43093	1.91530	0.70149	0.28784
2	0.45547	1.92480	0.71217	0.29372	12	0.43294	1.93440	0.70581	0.28606	22	0.48216	1.89670	0.69215	0.30095
3	0.50421	2.15480	0.81208	0.31640	13	0.52425	1.99240	0.73603	0.32146	23	0.42135	2.02730	0.75264	0.28210
4	0.48299	1.98320	0.73462	0.30725	14	0.42459	1.98790	0.73918	0.28454	24	0.41574	1.97580	0.73009	0.27768
5	0.48373	1.97170	0.72560	0.30851	15	0.50878	1.96940	0.71652	0.31052	25	0.41332	1.94210	0.72065	0.27555
6	0.49604	1.97910	0.73622	0.30974	16	0.50182	2.00090	0.73613	0.31401	26	0.38770	1.95250	0.72728	0.27222
7	0.48543	1.96850	0.72416	0.30495	17	0.45940	1.94890	0.71771	0.29682	27	0.46642	1.89110	0.69409	0.30313
8	0.43495	1.94540	0.71212	0.28537	18	0.52678	1.95450	0.70951	0.32021	28	0.45957	1.94960	0.72133	0.29477
9	0.45676	1.97620	0.73054	0.29730	19	0.46663	1.89430	0.70011	0.30437	29	0.48226	1.96640	0.72971	0.30787
10	0.44851	1.92350	0.71256	0.29480	20	0.46215	1.93260	0.71083	0.29829	30	0.39712	1.96770	0.73003	0.26961

Table-8.2: Descriptive Statistics of Estimated Parameters of Johnson S_U Distribution fitted to Thirty Sub-Samples (Size=5000) of JIF data for 2008								
Parameters	Median	Mean	Std. Error	Conf. -95%	Conf. 95%	Minimum	Maximum	Std. Dev.
γ	0.460860	0.460835	0.006596	0.447345	0.474325	0.387700	0.526780	0.036127
δ	1.960450	1.962353	0.009158	1.943624	1.981083	1.891100	2.154800	0.050159
λ	0.722745	0.724350	0.004100	0.715965	0.732735	0.692150	0.812080	0.022454
ζ	0.297060	0.296796	0.002590	0.291498	0.302094	0.269610	0.321460	0.014188

.

Fig.5: Temporal Changes in the Mean Estimated Parameters of Johnson SU Distribution of log10(JIF) 2001-2008



.

Fig.6: Confidence Interval about Mean Zeta 2001-2008

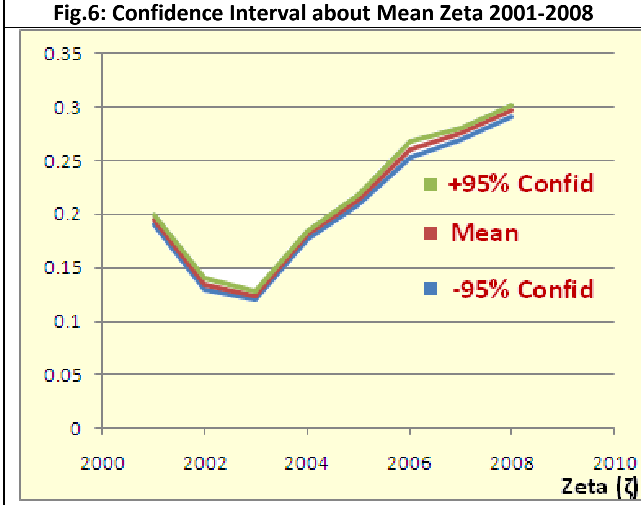
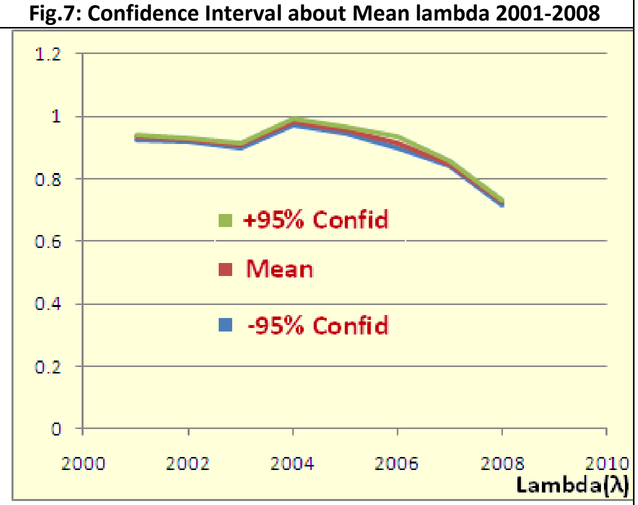


Fig.7: Confidence Interval about Mean lambda 2001-2008



.

Fig.8: Confidence Interval about Mean Delta 2001-2008

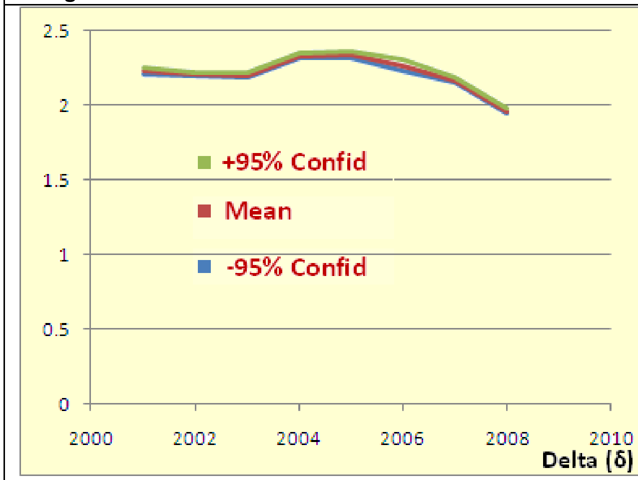
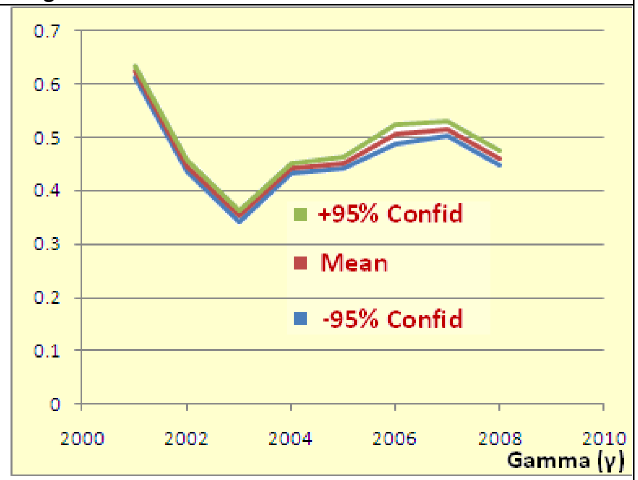


Fig.9: Confidence Interval about Mean Gamma 2001-2008



.

Table-9: Estimated Parameters of Pearson's Distribution for $\log_{10}$(JIF) Data for the Years 2001-2008

Year	N	b_0	b_1	b_2	Root of *f(x)*	Std Dev.	Skewness	Kurtosis	K	Type
2001	5679	-0.16830	0.06748	-0.09355	(0.36063, ±1.29186)	0.48369	0.44584	1.40520	0.07229	IV
2002	5475	-0.15426	0.04146	-0.09941	(0.20853, ±1.22811)	0.46884	0.29360	1.34073	0.02802	IV
2003	5702	-0.15394	0.03884	-0.09771	(0.19873, ±1.23932)	0.46667	0.27324	1.27974	0.02507	IV
2004	5913	-0.16249	0.04421	-0.08882	(0.24890, ±1.32951)	0.47066	0.29140	1.10631	0.03386	IV
2005	6033	-0.15427	0.04395	-0.08846	(0.24839, ±1.29701)	0.45130	0.29682	1.10484	0.03538	IV
2006	6152	-0.15156	0.05170	-0.09406	(0.27485, ±1.23925)	0.46925	0.36079	1.29544	0.04688	IV
2007	6226	-0.14351	0.05484	-0.09743	(0.28146, ±1.18057)	0.45041	0.39913	1.42420	0.05378	IV
2008	6545	-0.12715	0.05325	-0.10978	(0.24252, ±1.04855)	0.43542	0.43606	1.81472	0.05078	IV

K is Pearson's K for classification of empirical distributions obtained by solving $f(x) = b_0 + b_1 x + b_2 x^2 = 0$ (see Gupta and Kapoor, 1982; p 545)

5. Concluding Remarks: The results reported in this paper corroborate our earlier findings; the first that $\log_{10}$(JIF) is Pearson-IV distributed (Mishra, 2009), the second that although Burr and Dagum distributions fit very well to the data but they also exhibit instability of parameters over-the-samples, and the third that Johnson S_U distribution fits very well to the data and yields parameters stable over the samples (Mishra, 2010_b). Hence we conclude that Johnson S_U distribution is the best choice to fit to the $\log_{10}$(JIF) data. We have also found that over the years the $\log_{10}$(JIF) distribution is becoming more skewed and leptokurtic, possibly suggesting the Mathew effect (Tol, 2009) in operation, which means that more cited journals are cited ever more over time.

References

1. Draper, J. (1952) "Properties of Distributions Resulting from Certain Simple Transformations of the Normal Distribution", *Biometrika*, 39(3-4): 290-301.

2. George, F. (2007) *Johnson's System of Distributions and Microarray Data Analysis*, Doctoral Dissertation, Department of Mathematics, College of Arts and Sciences, University of South Florida, available at http://kong.lib.usf.edu:1801/view/action/singleViewer.do?dvs=1268014385536~52&locale=en_US&search_terms=000031367&application=DIGITOOL-3&frameId=1&usePid1=true&usePid2=true

3. Gupta, S.C. and Kapoor, V.K. (1982) Fundamentals of mathematical statistics [8th Edition], Sultan Chand & Sons, New Delhi.

4. Johnson, N.L.(1949) "Systems of Frequency Curves Generated by Methods of Translation", *Biometrika* 36(1-2): 149-176.

5. Mishra, S. K. (2009) "Does the Journal Impact Factor Help Make a Good Indicator of Academic Performance?", available at SSRN: http://ssrn.com/abstract=1485868

6. Mishra, S. K. (2010_a) "A Note on Empirical Sample Distribution of Journal Impact Factors in Major Discipline Groups", available at SSRN: http://ssrn.com/abstract=1552723

7. Mishra, S. K. (2010_b) "Empirical Probability Distribution of Journal Impact Factor and Over-the-Samples Stability in its Estimated Parameters", available at SSRN: http://ssrn.com/abstract=1556281

8. Slifker, J. and Shapiro, S. (1980) "The Johnson System: selection and parameter estimation", *Technometrics*, 22(2): 239-247.

9. Tadikamalla, P. R. (1980) "On Simulating Non-normal Distributions", *Psychometrica*, 45(2): 273-279.

10. Tol, R. S. J. (2009). "The Matthew effect defined and tested for the 100 most prolific economists", *Journal of the American Society for Information Science and Technology*, 60(2): 420-426.

11. Wheeler, R. (1980) "Quantile Estimators of Johnson curve Parameters", *Biometrika*, 67(3): 725-728.