A DISTRIBUTION OF PAPERS BASED ON FRACTIONAL COUNTING: AN EMPIRICAL STUDY

I.K. Ravichandra Rao, Professor and Head, Documentation Research and Training Centre, Indian Statistical Institute, 8th Mile Mysore Road, Bangalore 560059, India
e-mail: ikrrao@hotmail.com

Bibhuti Bhusan Sahoo, Junior Research Fellow, Documentation Research and Training Centre, Indian Statistical Institute, 8th Mile Mysore Road, Bangalore 560059, India
e-mail: bibhuti@isibang.ac.in and

L. Egghe, Professor, LUC, Universitaire Campus, B-3590 Diepenbeek, Belgium and UIA, Universiteitsplein 1, B – 2610 Wilrijk, Belgium. e-mail:leo.egghe@luc.ac.be

Abstract. *Distributions of papers based on the fractional counting are very irregular. It can be explained by a model which may be derived under the assumptions that the distribution of papers (φ (n))(method of normal counting) is a negative binomial distribution and the distribution of authors (ψ(n)) (multiple authorship) is a Poisson distribution. This model appears to be a much better model than the one which is derived earlier by Egghe and Rao under the assumption that φ (n) and ψ(n) confirm to Lotka's law.*

1. INTRODUCTION

Distributions of articles over authors are approximated by a number of related models since the first publication on the frequency distribution of scientific productivity by A J Lotka in 1926. The following are some of the important models or distributions, which are discussed since then:

1. Law of inverse square [9]
2. Generalized bibliometric distributions [1]
3. Negative binomial and as a special case, some times, geometric distribution [11]
4. Cumulative advantage distribution [10]

These distributions are related to a great extent – the geometric distribution is a special case of negative binomial distribution. The negative binomial distribution and also the cumulative advantage distribution may be derived from an urn model; they explain highly skewed data with a long tail fairly well. Further (1), (3) and (4) are used primarily to analyze size-frequency data. (1) is also a special case of (2) and they are discussed in a number of informetric papers. In most of these studies, the number of publications is considered as a measure of scientific productivity. As pointed out by Egghe [4] and Lindsey [8], there are three methods of counting the number of publications. They are:

I.K. Ravichandra Rao, Bibhuti Bhusan Sahoo and Leo Egghe, *A distribution of papers based on the fractional counting: an empirical study*, In: Guohua Jiang, Ronald Rousseau, Yishan Wu (Eds), Proceedings of the 9th International Conference on Scientometrics and Informetrics – ISSI 2003, Dalian University of Technology, Dalian, China, 2003, 241-250.

- Method of total counting or normal counting – assigning every author a weight one for each of his or her publications during a time period, irrespective of whether he or she is a first author or a second author, etc.
- Method of straight counting – assigning only the first author a weight one for each of his or her publications during a time period and for other authors a weight zero. In deriving the law of Inverse Square, Lotka adopted this method while collecting the data from Author Index of Chemical Abstracts and Auerbach's Geschichtstafeln der Physik.
- Method of fractional counting – assigning every author a weight 1/n in an n-authored paper.

Rousseau [13] in 1992 in his article entitled "Breakdown of the Robustness Property of Lotka's Law: The Case of Adjusted Counts for Multi-authorship Attribution", discussed frequency distribution of "fractional scores" in a bibliography of Informetrics. He observed that fractional counting of authors does not lead to a Lotka distribution. He further argued that Bookstein's robustness property of Lotka's law breaks down in such cases. Bookstein pointed out that "if we would find a $1/X^{\alpha}$ relation to describe productivity when we give a full publication to every author whose name appears on a paper, this will also be the case if we had assigned factional authorship instead" (2, p.383). This is one of the robustness properties. Ravichandra Rao [12] also studied a distribution of fractional scores in mathematics. His study was based on the bibliographical records in Math Reviews (1990.) For the appropriate groups or classes of fractional scores, he hypothesized that log-normal distribution fits much better than other distributions; however, this hypothesis was rejected when appropriate tests were applied. Recently, Egghe and Ravichandra Rao [5] further analyzed this data and came out with an extremely good model to describe distribution of fractional scores. Their paper entitled "Duality Revisited: Construction of Fractional Frequency Distributions based on two dual Lotka's Laws," is the first attempt of this kind. They have assumed two simple Lotka distributions with exponent 2 – one for the number of authors with n papers (total count) and the other one for the number of papers with n authors. Based on the earlier convolution model of Egghe [4], the authors have reworked for discrete scores and produced a theoretical fractional frequency distribution (f(q)) with only one parameter which is in very close agreement with observed data, produced earlier by Rao. Egghe and Ravichandra Rao thus concluded that "fractional distributions are a consequence of Lotka's law and are not examples of breakdowns of this famous historical law." Further, they have also noticed that a Poisson distribution (for $\psi(n)$) if the parameter λ is chosen in the appropriate way is better capable of describing the distribution of fractional scores (the results have not been published). Thus, as a continuation of Egghe and Ravichandra Rao's work, an attempt has been made here to

- Identify a suitable model for distribution of papers in the field of software studies, as we find in many cases Lotka's law hardly fits
- Identify a suitable model to describe distribution of authors (distribution of multiple authorship!) and then
- Identify an appropriate model to explain the distribution of fractional scores of authors.

2. DATA COLLECTION

Data in the area of "software and related topics" were collected from the COMPENDEX database for the year 2000. After eliminating duplicate records, there were a total of 55,784 relevant records. All the three methods – total counting, straight counting and fractional counting – were adopted to collect the data on distribution of papers over authors. Further, data on distribution of authors (multiple authorship) over papers were also collected. The data are given in Table 1. Table 2 gives the distribution of papers, based on fractional counting.

3. DATA ANALYSIS

Lotka observed regularities in the productivity of chemists and physicists and on the basis of these observations, he formulated a hypothesis that the relative frequency of authors publishing x articles could be explained as

$$y = \frac{6}{\pi^2} x^{-\alpha}$$

where α is a constant. The value of α was found to be 2 for physicists and 1.89 for chemists. Since then, several formal analytical and predictive models have been developed for describing the phenomenon of scientific productivity [11]. Ravichandra Rao [11] in his article on distribution of scientific productivity and social change argued that the negative binomial distribution:

$$p(x) = \frac{(k+x-2)!}{(k-1)!(x-1)!} p^k q^x \qquad x = 1,2,3, \ldots$$

$$0 \leq p, q \leq 1, \ k > 0$$

fits fairly well to the author productivity data. Even in the present study, it has been observed that the negative binomial fits (data on author productivity) much better than most other distributions, such as Lotka's distribution, Poisson, lognormal, logarithmic series, geometric, etc. The negative binomial distribution, on fitting, gives a minimum chi-square value. The results of fitting the negative binomial distribution is shown in Table 1 (Total counting). Further, the authors have also observed in this study that the negative binomial distribution fits better than most of the other well-known distributions, to the data on author productivity, based on straight counting. Table 1 shows the results. An attempt has also been made to identify a suitable distribution to the distribution of multiple authorship. It has been observed that the Poisson distribution (x is modified such that x = 1,2,3,….)

$$p(x) = \frac{e^{-\lambda} . \lambda^{x-1}}{(x-1)!}, \qquad x = 1,2,3,4......$$

fits much better than any other well known probability distribution. Table 1 gives the results.

4. DISTRIBUTION OF FRACTIONAL PAPERS

Egghe and Ravichandra Rao [5] have derived a theoretical model for the fractional frequency distribution f(q) (discrete case) from two dual Lotka's laws. They have derived the required formula for f(q), q > 0, for different cases:

1) case 1: i = 2, allowing an author score of 1/2 or 1 in one paper
2) case 2: i =3, allowing an author score of 1/3,1/2 or 1 in one paper
3) case 3: i =4, allowing an author score of 1/4, 1/3,1/2 or 1 in one paper
4) case 4: i = 5, allowing an author score of 1/5 1/4, 1/3,1/2 or 1 in one paper

For each case i, fractional scores of 1/j for j larger than or greater than i are set to be 1/i for reasons of simplicity. The case 1 for i = 2 is the most simple one and the relevant formulae are:

$f(1/2) = g_1(1/2)\,\phi\,(1) = (1-f_1(1))\,\phi\,(1)$
$f(1) = g_1(1)\,\phi\,(1) + (\,g_1(1/2))^2\,\phi\,(2) = f_1(1)\,\phi\,(1) + (1-f_1(1))^2\,\phi\,(2)$
$f(3/2) = 2g_1(1/2)\,g_1(1)\,\phi\,(2) + (\,g_1(1/2))^3\,\phi\,(3) = 2(1-f_1(1))\,f_1(1)\,\phi(2) + (1-f_1(1))^3\,\phi\,(3)$
$f(2) = (g_1(1))^2\,\phi\,(2) + 3(\,g_1(1/2))^2\,\,g_1(1)\,\phi\,(3) + (\,g_1(1/2))^4\,\phi\,(4)$
$\quad = (f_1(1))^2\,\phi(2) + 3(1-\,f_1(1))^2\,f_1(1)\,\phi\,(3)\,\,+ (1-f_1(1))^4\,\phi\,(4)$

$g_1(.)$ is derived using $f_1(.)$ and ϕ (.) in each of the cases discussed above as in Egghe and Rao[4]. Formulae for cases i = 3 & 4 are given by Egghe and Ravichandra Rao [5]. Formulae for case i = 5 are too many and run into several pages and therefore they are not published so far. They are however available with the authors, if required. In the formula for f(q), ϕ (n) is the distribution of papers over authors (Lotka's law), and

$$f_1(z) = \frac{\Psi\left(\frac{1}{z}\right)}{\mu z}$$

where ψ(z) is the distribution of papers with z authors; μ denotes the average number of authors per paper; thus $f_1(z)$ denotes the fraction of authors with fractional score z in one paper. Using the above formula, Egghe and Ravichandra Rao [5] under the assumption that both ψ(n) and ϕ (n) confirm to a Lotka's law, computed probabilities ($f_1(q)$) for cases i = 2,3,4 and 5. As noted in their article the results were excellent, particularly (for the case i = 5.)

In this paper a similar attempt is made and under the assumption that both ψ(n) and ϕ (n) confirm to a Lotka's law, to compute f(q) and it is not giving a good result (Table 3, Cols 4 & 8.) As may be observed from Table 1, both ψ(n) and ϕ (n) do not confirm to a Lotka's Law and this may be a reason for the bad result of f(q). On the other hand, we have observed that Poisson distribution is a close approximation to ψ(n). Therefore an attempt was made to compute f(q), under the assumption that ψ(n) follows a Poisson distribution and ϕ (n) confirms to Lotka's Law; the results are given in Tables 3 (Cols 4 & 9) for case i=4. Figures 1-4 give the graphs for theoretical and empirical values for cases i = 2 to 5. However, as may be observed in Figures 1-4, the results are not satisfactory. The moments method is used to estimate the parameters, while fitting the distribution.

Ravichandra Rao [11] argued that the negative binomial distribution describes a pattern of scientific productivity under the success breeds-success condition in a wide variety of social changes. Further even in the present study, authors have observed that the negative binomial distribution fits ϕ(n) fairly well. Therefore, an attempt has been made here to compute f(q) for cases i = 2,3,4 and 5 under the assumption that ψ(n) and ϕ (n) follow Poisson and negative binomial distributions respectively. The values of f(q) are very close to the experimental values and the results are excellent. This observation is based on Figures 1-4. The theoretical and observed values for case i = 4 is given in Table 3 (Cols 5 & 10.) Figures 1-4 give the graphs for theoretical and empirical values for cases i = 2 to 5. The $g_1(.)$ were accordingly derived for cases i =2 to 5 and they are as follows:

Case 1: i=2

This is a simple case. In this case, an author receives a score 1, if he / she is an author in a single authored paper. If he/she is an author in an multi-authored paper, the author receives a score 1/2.

$$g_1(1) = f_1(1) = e^{-\lambda} / (\lambda + 1) \text{ and } g_1(1/2) = 1-f_1(1)$$

Case 2: i=3

In this case, an author receives a score 1, if he / she is an author in a single authored paper. The author receives a score of ½ if he / she is an author in a two-authored paper. A score of 1/3 is assigned if he / she is an author in a j-authored paper for all j ≥ 3.

$$g_1(1) = f_1(1) = e^{-\lambda} / (\lambda + 1),\ g_1(1/2) = 2 \lambda\, g_1(1) \text{ and } g_1(1/3) = 1- (f_1(1) + f_1(1/2))$$
$$= 1- e^{-\lambda} / (\lambda + 1)\{1+2\lambda\} = 1- g_1(1)\,(1+2\lambda)$$

Case 3: i=4

In this case, an author receives a score 1, if he / she is an author in a single authored paper. The author receives a score of ½ if he / she is an author in a two-authored paper. A score of 1/3 is assigned if he / she is an author in a 3-authored paper and a score of ¼ is assigned if he / she is a j-authored paper for all j ≥ 4.

$$g_1(1) = f_1(1) = e^{-\lambda} / (\lambda + 1),\ g_1(1/2) = 2 \lambda\, g_1(1),\ g_1(1/3) = 3\,(\lambda/2)\, g_1(1/2)$$
$$= 1.5\lambda^2 g_1(1) \text{ and } g_1(1/4) =1-g_1(1)\,(1+2\lambda+1.5\lambda^2)$$

Case 4: i=5

As in cases 1,2 and 3 the author receives a score 1/j, if he / she is an author in a j-authored paper (j ≤ 5) and the author receives a score of 1/5 if he / she is an author in a j-authored paper (j ≥ 5).

$$g_1(1) = f_1(1) = e^{-\lambda} / (\lambda + 1),\ g_1(1/2) = 2 \lambda\, g_1(1),\ g_1(1/3) =1.5\lambda^2 g_1(1),\ g_1(1/4)$$
$$= (\lambda^2/3)\, g_1(1/2) = (2/3)\, \lambda^3 g_1(1) \text{ and } g_1(1/5) =1-g_1(1)\,(1+2\lambda+1.5\lambda^2+(2/3)\, \lambda^3)$$

In all the above cases (i = 2,3,4 and 5), fractional scores of 1/j for j larger than or equal to i , are set to be 1/i for reasons of manageability of the calculations. The larger i, the better the scoring system. Also $g_1(1)$ is derived using $f_1(1)$. $g_1(1/2)$, $g_1(1/3)$, $g_1(1/4)$ and $g_1(1/5)$ are functions of $g_1(1)$. $g_1(.)$ refers to the author distribution of fractional scores in one paper.

5 . CONCLUSION

In a working hypothesis that the population is a mixture of individuals with different degrees of accident proneness, represented by different λ in a Poisson distribution and if suppose that in the population the distribution of λ is of the Gamma form [6,7], then the variable X follows a negative binomial distribution. In the case of author productivity, each individual author has different capabilities to publish an article (represented by different λ -- similar to that of accident proneness.) Further, as has been observed in the literature earlier [5], and as observed in this article (based on total counting), it has been hypothesized that the distribution of papers over authors confirm to a negative binomial distribution. Since $\psi(n)$ closely confirms to a Poisson distribution and $\phi(n)$ confirms to a negative binomial distribution (-- a compound Poisson distribution), in this paper it is further conjectured that f(q) belongs to a family of Poisson distribution and it explains the scientific productivity of author, to a great extent. The values of Kolmogorov statistics (D_{max}) for Lotka-Lotka, Poisson-Lotka and Neg. Bin.- Poisson cases are 0.330204, 0.26399, and 0.019059 respectively. It undoubtedly leads us to conclude that f(q) can reasonably be predicted under the assumption that $\varphi(n)$ follows Negative binomial and $\psi(n)$ follows Poisson distribution .

REFERENCES

1) Bookstein, A. Patterns of Scientific Productivity and Social Change: A discussion of Lotka's Law and Bibliometric Symmetry. (Journal of the American Society for Information Science, 28(4), 1977, 206-10)
2) Bookstein, A. Informetric Distributions, Part II: Resilience to Ambiguity. (Journal of the American Society for Information Science, 41(5), 1990, 376-386)
3) Egghe, L. The Duality of Informetric Systems with Applications to the Empirical Laws. (Journal of Information Science, 16(1), 17-27.) (Also *see* Introduction to Informetrics : Quantitative Methods in Library, Documentation and Information Science by L Egghe and R Rousseau. Elsevier. Amsterdam. 1990.)
4) Egghe, L. Consequence of Lotka's law in the case of Fractional counting of authorship and of Author Count. (Math Computing Modeling. 18(9), 1993, 63-77)
5) Egghe, L and Rao, I.K. Ravichandra. Duality Revisited: Construction Fractional Frequency Distributions based on two dual Lotka Laws (Journal of the American Society for Information Science and Technology; 53(10), 2002, 789-801)
6) Johnson, L Norman and Kotz, Samuel. Distributions in Statistics: Discrete Distributions. Houghton Mifflin Co. Boston. 1969.
7) Kendall, Maurice and Stuart, Alan. The Advanced Theory of Statistics. Vol. 1. Distribution Theory. E. 4. 1977. p. 136-7.
8) Lindsey, D. Production and Citation Measures in the Society of Science: The Problem of Multiple Authorship (Social Studies of Science, 10, 1980, 145-62)
9) Lotka, A J. The Frequency Distribution of Scientific Productivity. (Journal of the Washington Academy of Science. 16(12), 1926, 317-23)
10) Price, D de Solla. A General Theory of Bibliometric and other Cumulative Advantage Processes. (Journal of the American Society for Information Science. 27(5), 1976, 292-306)
11) Rao, I.K.Ravichandra. The Distribution of Scientific Productivity and Social Change. (J of the American Society for Information Science., 1980, 111-22)

12) Rao, I.K.Ravichandra. A stochastic Approach to Analysis of Distribution of Papers in Mathematics: Lotka's Law Revisited.(Proceedings of the Fifth Biennial Conference of the International Society for Scientometrics and Informetrics. Royal College, River Forest, IL; June 7-10, 1995. p. 455-64)

13) Rousseau, R. Breakdown of the Robustness of Lotka's Law: The case of Adjusted Counts for Multi-authorship Attribution. (Journal of the American Society for Information Science; 43(10), 1992, 645-47)

Table 1. Author Productivity in the area of Software Studies

No. of Papers	No. of Authors (Total Counting)	Theoretical Values (neg. binomial)	No. of Authors (Straight Counting)	Theoretical Values (neg. binomial)	No. of Authors	No. of Papers	Theoretical Values (Poisson)
N	$\varphi(n)$		$\varphi(n)$		*y*	$\psi(n)$	
1	110503	111787	41071	41451	1	8820	8940
2	14027	11831	4742	4081	2	18555	16369
3	3223	3665	1012	1126	3	14577	14985
4	1085	1385	254	378	4	7536	9146
5	414	570	96	138	5	3320	4186
6	195	246	15	53	6	1429	1533
7	110	110	14	21	7	681	468
8	45	50	7	8	8	356	122
9	26	23	1	3	9	211	28
10	15	11	4	1	10	118	6
11	17	5	2	1	11	68	1
12	9	2	2	0	12	37	
13	8	1	3		13	30	
14	4	1	-		14	25	
15	2		-		15	21	
16	2		-				
17	1		-				
18	1		-				
22	1		-				
24	1		1				
Total	129689		47268			55784	
Mean	1.2177		1.1802			2.8309	
Variance	0.4482		0.3297			2.4288	
St. dev.	0.6695		0.5742			1.5585	
p		0.4862		0.5464			-
q		0.5138		0.4536			-
k		0.2060		0.2170			-
λ		-		-			1.8309
D_{max}		0.0036		0.0082			0.0370
D_{α}		0.0038		0.0063			0.0058

Table 2. Distribution of Papers (Fractional Method)

Fraction of Papers	No. of Authors
Z	f(z)
0.0667	219
0.0714	250
0.0769	284
0.0833	333
0.0909	586
0.1	875
0.1111	1377
0.125	2121
0.1334	12
0.1381	14
0.1428	1
0.1429	3396
0.1436	7
0.1483	2
0.15	2
Fraction of Papers	No. of Authors
Z	f(z)
0.1538	7
0.1547	8
0.1623	1
0.1667	6144
0.1678	7
0.1714	9
0.1742	2
0.1769	9
0.1778	8
0.1818	3
0.1825	8
0.1909	12
0.1917	2
0.1944	11
0.1964	3
Fraction of Papers	No. of Authors
Z	f(z)
0.2	11555
0.2019	7
0.202	5

0.2083	8
0.2096	3
0.2111	8
0.2143	9
0.2159	4
0.2198	2
0.2222	41
0.225	8
0.2269	1
0.2334	3
0.2338	9
0.2361	30
Fraction of Papers Z	No. of Authors f(z)
0.2429	9
0.2436	6
0.2456	2
0.25	22197
0.254	44
0.2667	21
0.2679	58
0.2763	1
0.2769	10
0.2778	49
0.2833	8
0.2858	116
0.2909	17
0.2917	65
0.2967	3
Fraction of Papers Z	No. of Authors f(z)
0.2976	4
0.3	26
0.3028	1
0.3096	152
0.3103	2
0.3107	2
0.3111	40
0.3194	3
0.325	80
0.3269	1
0.3333	29695
0.3334	213
0.3338	5
0.3409	1
0.3429	122
Fraction of Papers Z	No. of Authors f(z)
0.3445	3

0.35	2
0.3512	1
0.3572	5
0.3576	1
0.3651	1
0.3667	363
0.3678	1
0.3767	2
0.379	3
0.3917	2
0.3929	8
0.3936	2
0.3969	2
0.4005	1
0.402	5
0.4028	1
0.4047	4
0.4083	1
0.4102	6
0.4108	8
0.4166	17
0.4167	3
0.4207	1
0.4222	1
0.4242	24
0.425	1
0.4287	3
0.4333	40
0.4334	4
0.4338	2
0.4346	5
0.4361	1
0.4444	65
0.4445	5
0.45	6
0.4525	11
0.454	2
0.4576	2
0.4583	105
0.4584	8
0.4667	3
0.4679	7
0.4714	2
0.4722	1
0.4762	158
0.4763	16
0.4778	6
0.4858	8
0.4901	1
0.4909	1
Fraction of Papers Z	No. of Authors f(z)
0.4917	17

0.5	27259
0.5001	25
0.5011	1
0.5075	1
0.5111	3
0.5167	1
0.5222	1
0.5242	1
0.525	15
0.5255	1
0.5277	2
0.5333	896
0.5334	41
0.5429	21
0.5444	1
0.5448	1
0.5525	3
0.5531	1
0.5537	2
0.5555	3
0.5596	1
0.5603	1
0.5636	1
0.5667	44
0.5671	1
0.5694	3
0.5714	1
0.5716	1
0.5762	1
0.5775	3
0.5833	2
0.5858	2
0.5873	6
0.5874	1
0.5909	1
0.5969	1
0.6	73
0.6012	1
0.6013	1
0.602	1
0.6102	1
0.6108	1
0.6111	6
0.6191	12
0.6192	1
0.6198	1
0.6222	1
0.6242	4
0.6243	1
0.625	13
Fraction of Papers Z	No. of Authors f(z)
0.6251	1

0.6262	1
0.63	1
0.6334	2
0.6338	1
0.6346	1
0.6429	14
0.643	5
0.6444	7
0.6445	1
0.6449	1
0.65	1
0.6525	2
0.6531	1
0.6583	10
0.6648	1
0.6666	3057
0.6667	25
0.6668	2
0.6679	4
0.6762	22
0.6763	6
0.6778	4
0.6917	3
0.6922	1
0.6944	1
0.7	62
0.7001	2
0.7025	1
0.7096	1
0.7111	1
0.7333	108
0.7334	4
0.7353	2
0.7401	1
0.7429	5
0.7499	4
0.75	1713
0.7525	3
0.7575	2
0.762	2
0.7666	4
0.7667	4
0.7679	3
0.7714	1
0.7762	1
0.7777	10
0.7778	1
0.7833	1
0.7858	3
0.7916	11
Fraction of Papers Z	No. of Authors f(z)
0.8	17

0.8095	40
0.8096	3
0.8131	1
0.8192	1
0.8194	1
0.825	2
0.83	1
0.8333	64
0.8339	1
0.8346	3
0.8353	1
0.8429	5
0.843	1
0.8444	3
0.8525	1
0.8583	5
0.8666	183
0.8667	7
0.8762	3
0.8825	1
0.8858	1
0.887	3
0.8928	2
0.8936	1
0.9	13
0.9001	1
0.9095	1
0.9166	2
0.9198	1
0.9262	1
0.9333	10
0.9345	1
0.9361	1
0.9435	1
0.9445	1
0.9524	2
0.9583	3
0.9666	3
0.9667	1
0.9762	3
0.9777	2
0.9916	1
0.9917	1
0.9999	581
1	9318
1.0095	2
1.0111	1
1.0263	1
1.0333	14
1.0334	1
Fraction of Papers Z	No. of Authors f(z)
1.0429	1

1.043	1	1.2917	1	1.5666	2	2.0334	1	2.7333	1
1.0493	1	1.3	1	1.5715	1	2.0665	4	2.75	25
1.0583	1	1.3095	1	1.5789	1	2.0666	2	2.7667	2
1.0584	1	1.3096	2	1.5999	4	2.0667	1	2.8332	1
1.0666	36	1.3194	1	1.6	3	2.125	1	2.8664	2
1.0667	1	1.3249	2	1.6095	1	2.1333	1	2.8666	1
1.0714	1	1.325	1	1.611	1	2.1428	1	2.9094	1
1.0769	1	1.3332	164	1.6111	1	2.1429	1	2.9666	1
1.0873	1	1.3333	428	1.6167	1	2.1582	1	2.9667	2
1.0909	6	1.3334	2	1.625	1	2.1665	4	2.976	1
1.0953	1	1.3372	1	1.6333	2	2.1666	2	2.9998	1
1.0999	1	1.3428	2	1.6665	62	2.1667	3	2.9999	2
1.1	8	1.3429	4	1.6666	75	2.1679	1	3	105
1.1012	1	1.3666	2	1.6667	1	2.1777	1	3.0343	1
1.1028	1	1.3667	6	1.6762	1	2.1916	1	3.0666	1
1.111	1	1.3858	1	1.6999	3	2.1998	7	3.0833	1
1.1111	10	1.3916	1	1.7	1	2.1999	3	3.0909	1
1.1191	1	1.3999	12	1.7332	7	2.2	5	3.1012	1
1.1192	1	1.4	9	1.7333	5	2.25	49	3.1429	1
1.1242	1	1.4012	1	1.738	1	2.2679	1	3.1997	4
1.1249	2	1.4019	1	1.75	113	2.2917	1	3.1998	1
1.125	9	1.4077	1	1.7666	1	2.311	1	3.2	2
1.1333	4	1.4095	1	1.7776	1	2.3111	1	3.25	9
1.1361	1	1.4123	1	1.7777	1	2.3331	10	3.2664	1
1.1428	7	1.4241	1	1.7789	1	2.3332	9	3.333	1
1.1429	22	1.4242	1	1.8094	1	2.3333	42	3.3331	1
1.1525	1	1.4334	2	1.8095	3	2.3998	3	3.3332	1
1.1666	16	1.4443	1	1.8249	1	2.3999	2	3.3333	7
1.1667	36	1.4444	1	1.8333	2	2.4	2	3.4583	1
1.1762	1	1.4582	1	1.8428	1	2.4242	1	3.5	21
1.1777	2	1.4583	4	1.8582	1	2.4581	1	3.533	1
1.188	1	1.4666	3	1.8665	16	2.4678	1	3.5331	1
1.1916	3	1.4667	1	1.8666	10	2.4762	1	3.5333	1
1.1999	53	1.4761	5	1.9	1	2.4999	1	3.5998	1
1.2	82	1.4762	3	1.9332	2	2.5	85	3.6663	1
1.2096	1	1.4778	1	1.9691	1	2.5331	1	3.6666	3
1.2221	1	1.4857	2	1.9762	1	2.5332	2	3.733	1
1.225	1	1.4999	8	1.9776	1	2.6	1	3.75	7
1.2251	1	1.5	677	1.9998	20	2.6664	4	3.8663	1
1.2333	5	1.5095	1	1.9999	27	2.6665	2	3.8759	1
1.2382	1	1.5249	1	2	604	2.6666	10	3.9999	3
1.2435	1	1.5332	17	2.0095	1	2.676	1	4	30
1.2499	1								
1.25	344								
1.2539	1								
1.2611	1								
1.2666	4								
1.2667	3								
1.2713	1								
1.2762	1								

Fraction of Papers	No. of Authors	Fraction of Papers	No. of Authors	Fraction of Papers	No. of Authors	Fraction of Papers	No. of Authors	Fraction of Papers	No. of Authors
Z	f(z)	Z	f(z)	Z	f(z)	Z	f(z)	Z	f(z)
1.2909	1	1.5333	18	2.0332	5	2.7331	1	4.0094	1

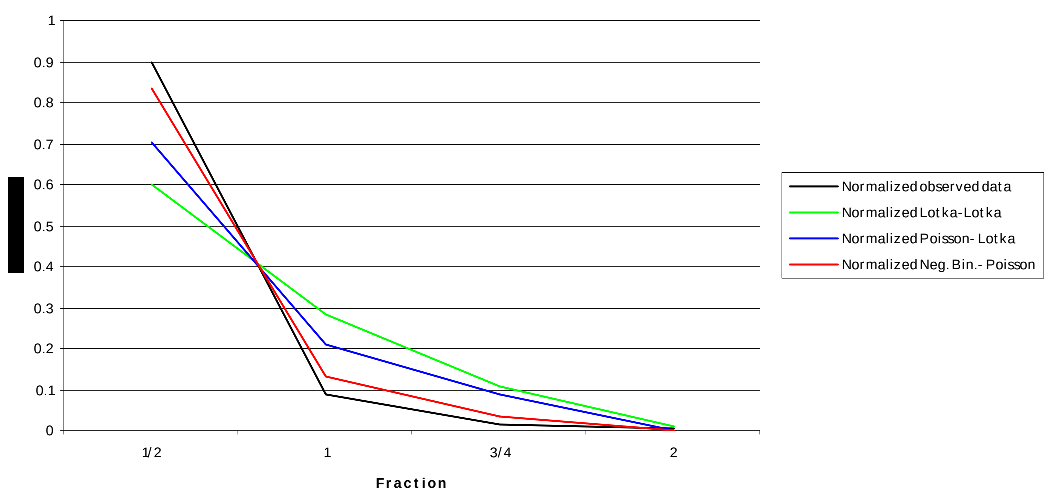
4.25	6	5.75	2	13	2
4.3333	2	6	5		
4.3667	1	6.5	1		
4.5	9	6.5327	1		
4.5331	1	7	6		
4.5332	1	8	4		
4.6666	1	8.3333	1		
5	15	9	4		
5.25	3	10	2		
5.2759	1	11	1		
5.5	4	12	2		

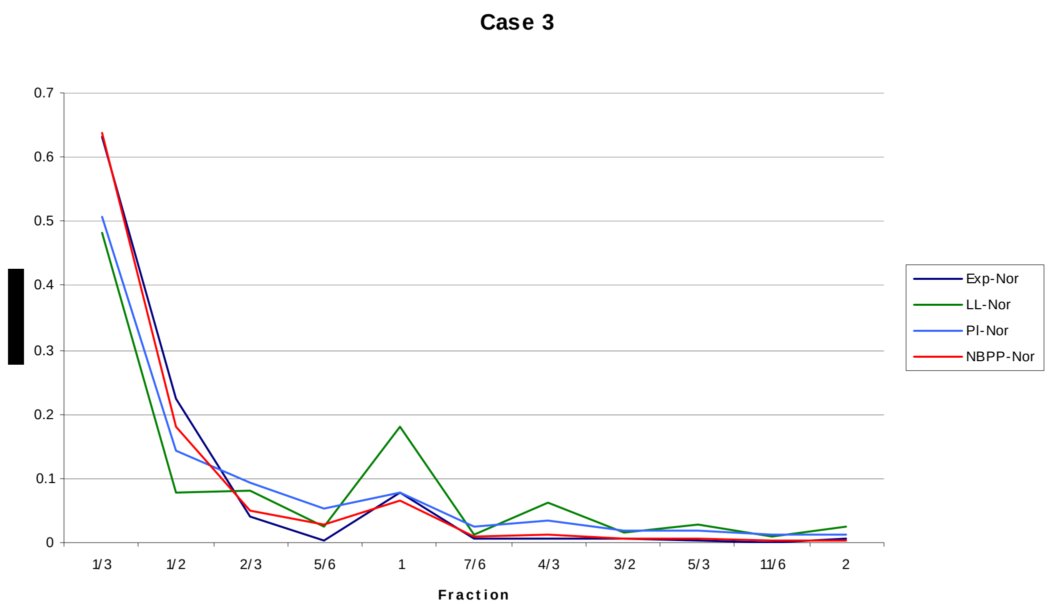
Table 3. Values of f(q) for case i = 4

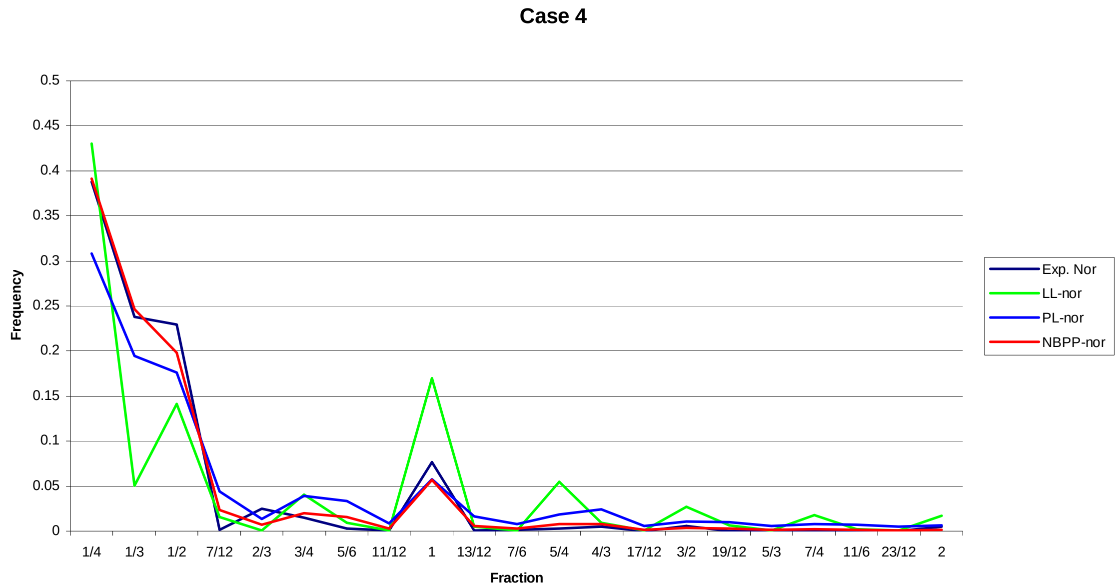
q	Obs. Data f(q)	Lotka-Lotka	Poisson-Lotka	Neg. Bin.-Poisson	q	Obs. Data f(q)	Lotka-Lotka	Poisson-Lotka	Neg. Bin.-Poisson
1	2	3	4	5	6	7	8	9	10
1/4	0.385931	0.368761	0.274408	0.389119	5/4	0.002853	0.046649	0.016635	0.007646
1/3	0.236928	0.04348	0.173053	0.245394	4/3	0.004758	0.008132	0.02121	0.007721
1/2	0.227992	0.121144	0.156991	0.197279	17/12	0.00027	0.000429	0.00527	0.001301
7/12	0.001689	0.013188	0.039058	0.023426	3/2	0.005714	0.023194	0.009779	0.003478
2/3	0.025014	0.000777	0.012316	0.007387	19/12	0.000131	0.005431	0.009007	0.002546
3/4	0.014535	0.034859	0.034657	0.019659	5/3	0.001118	0.000559	0.005119	0.001169
5/6	0.002637	0.007666	0.029692	0.015676	7/4	0.001002	0.014984	0.006718	0.001785
11/12	0.000378	0.000629	0.007413	0.003101	11/6	0.00027	0.001834	0.006214	0.001424
1	0.076568	0.145334	0.051085	0.056745	23/12	2.31E-05	0.000563	0.004138	0.000687
13/12	0.000679	0.004313	0.014776	0.005636	2	0.005097	0.014803	0.005747	0.001373
7/6	0.001766	0.00054	0.007169	0.002483	Total	0.99535	0.85727	0.890454	0.995034

Note: In Table 3, Lotka-Lotka means, both ψ(n) and φ(n) follow Lotka's law. Poisson-Lotka means, ψ(n) and φ(n) follows Poisson and Lotka's distributions respectively. Neg. bin– Poisson means φ(n) and ψ(n) follow Negative binomial and Poisson distributions respectively.

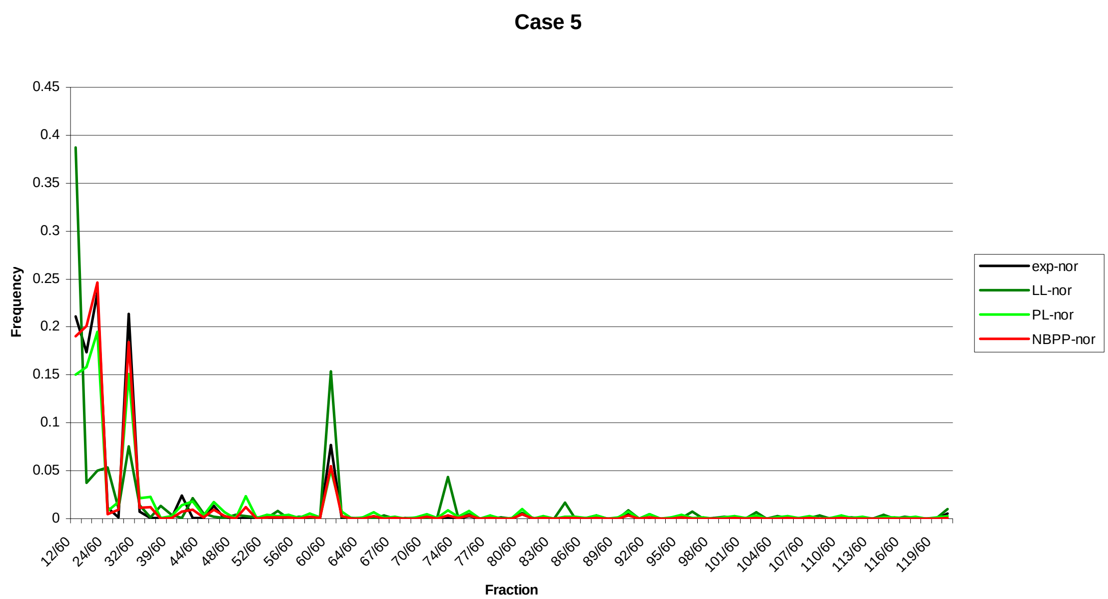
case 2




Case 3
[redacted]
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1/3
1/2
2/3
5/6
1
7/6
4/3
3/2
5/3
11/6
2
Fraction
Exp-Nor
LL-Nor
PI-Nor
NBPP-Nor



Case 4
Frequency
Fraction
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
1/4
1/3
1/2
7/12
2/3
3/4
5/6
11/12
1
13/12
7/6
5/4
4/3
17/12
3/2
19/12
5/3
7/4
11/6
23/12
2
Exp. Nor
LL-nor
PL-nor
NBPP-nor



Case 5
Frequency
Fraction
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
12/60
24/60
32/60
39/60
44/60
48/60
52/60
56/60
60/60
64/60
67/60
70/60
74/60
77/60
80/60
83/60
86/60
89/60
92/60
95/60
98/60
101/60
104/60
107/60
110/60
113/60
116/60
119/60
exp-nor
LL-nor
PL-nor
NBPP-nor