A DISTRIBUTION OF PAPERS BASED ON THE 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
e-mail:leo.egghe@luc.ac.be

[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 straight 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:

- Law of inverse square [8]
- Generalized bibliometric distributions [1]
- Negative binomial and as a special case, some times, geometric distribution [10]
- Cumulative advantage distribution [9]

In most of these studies, the number of publications is considered as a measure of scientific productivity. As pointed out by Egghe [3] and Lindsey [7], there are three methods of counting the number of publications. They are:

- 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 Aurbach's Geschichtstafeln der Physik.
- Method of fractional counting – assigning every author a weight 1/n in an n-authored paper.

Rousseau [12] 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. Ravichandra Rao [11] also studied a distribution of fractional scores in mathematics. His study was based on the articles covered in Math Reviews (1990.) For the appropriate groups or classes of fractional scores, he observed that lognormal distribution fits much better than other distributions; however, this hypothesis was rejected when appropriate tests were applied. Recently, Egghe and Ravichandra Rao [4] 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 [3], 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 [10]. Ravichandra Rao [10] in his article on distribution of scientific productivity and social change argued that the negative binomial distribution:

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

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

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 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.

4.1 Distribution of Fractional Papers

Egghe and Ravichandra Rao [4] 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, ½ or 1 in one paper

The case 1 for i = 2 is the most simple one and the relevant formulae are:

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

Formulae for cases i = 3 & 4 are given by Egghe and Ravichandra Rao [4]. Formula 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 (Loatka's law), and

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

where $\psi(z)$ 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 [4] under the assumption that both $\psi(n)$ and ϕ (n) confirm to a Lotka's law, computed probabilities ($f_i(q)$) for cases i = 1,2,3,4 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 $\psi(n)$ and ϕ (n) confirm to a Lotka's law, to compute f(q) and it is not giving a good result. As may be observed from Table 1, both $\psi(n)$ 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 $\psi(n)$. Therefore an attempt were made to compute f(q), under the assumption that $\psi(n)$ follows a Poisson distribution and ϕ (n) confirms to Lotka's Law; the results are given in Tables 3-6 and in Figures 1-4. However, as may be observed in Figures 1-4, the results are not satisfactory

Ravichandra Rao [10] 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 $\phi(n)$ fairly well. Therefore, an attempt has been made here to compute f(q) for cases i = 1,2,3,4 and 5 under the assumption that $\psi(n)$ 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. The results are shown Tables 3-6 and in Figures 1-4. The $g_1(.)$ were then accordingly derived 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.

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

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 $j \geq 3$.

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

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 $j \geq 4$.

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

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 ($j \leq 5$) and the author receives a score of 1/5 if he / she is an author in a j-authored paper ($j \geq 5$).

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

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.

4. 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 [5,6], 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 (similar to that of accident proneness.) Further, as has been observed in the literature earlier [4], 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.

6. BIBLIOGRAPHICAL 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) Egghe, L. The Duality of Informetric Systems with Applications to the Empirical Laws. (Journal of Information Science, 16(1), 17-27.)

3) Egghe, L. Consequence of Lotka's law in the case of Fractional counting of authorship and of Author Count. (Math Computing Modeling. Vol. 18(9), 1993, 63-77)

4) Egghe, L and Rao, I.K. Ravichandra. Duality Revisited: Construction Fractional Frequency Distributions based on two dual Lotka Laws (to be published in JASIST)

5) Johnson, L Norman and Kotz, Samuel. Distributions in Statistics: Discrete Distributions. Houghton Mifflin Co. Boston. 1969.

6) Kendall, Maurice and Staurt, Alan. The Advanced Theory of Statistics. Vol. 1. Distribution Theory. E. 4. 1977. p. 136-7.

7) Lindsey, D. Production and Citation Measures in the Society of Science: The Problem of Multiple Authorship (Social Studies of Science, 10, 1980, 145-62)

8) Lotka, A J. The Frequency Distribution of Scientific Productivity. (Journal of the Washington Academy of Science. 16(12), 1926, 317-23)

9) 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)

10) Rao, I.K.Ravichandra. The Distribution of Scientific Productivity and Social Change. (J of the American Society for Information Science., 1980, 111-22)

11) 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)

12) 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 (Strait Counting)	Theoretical Values (neg. binomial)	No. of Authors	No. of Papers	Theoretical Values (Poisson)
X	$f(x)$		$g(x)$		y	$f(y)$	
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	Cum. Freq.
z	f(z)	F(z)
0.0667	219	219
0.0714	250	469
0.0769	284	753
0.0833	333	1086
0.0909	586	1672
0.1	875	2547
0.1111	1377	3924
0.125	2121	6045
0.1334	12	6057
0.1381	14	6071
0.1428	1	6072
0.1429	3396	9468
0.1436	7	9475
0.1483	2	9477
0.15	2	9479
0.1538	7	9486
0.1547	8	9494
0.1623	1	9495
0.1667	6144	15639
0.1678	7	15646
0.1714	9	15655
0.1742	2	15657
0.1769	9	15666
0.1778	8	15674
0.1818	3	15677
0.1825	8	15685
0.1909	12	15699
0.1917	2	15701
0.1944	11	15712
0.1964	3	15715
0.2	11555	27270
0.2019	7	27277
0.202	5	27282
0.2083	8	27290
0.2096	3	27293
0.2111	8	27301
0.2143	9	27310
0.2159	4	27314
0.2198	2	27316
0.2222	41	27357
0.225	8	27365
0.2269	1	27377
0.2334	3	27380
0.2338	9	27389
0.2361	30	27419
0.2429	9	27435
0.2436	6	27441
0.2456	2	27443
0.25	22197	49640
0.254	44	49684
0.2667	21	49727
0.2679	58	49785
0.2763	1	49786
0.2769	10	49796
0.2778	49	49845
0.2833	8	49853
0.2858	116	49969
0.2909	17	49986
0.2917	65	50051
0.2967	3	50054
0.2976	4	50058
0.3	26	50084
0.3028	1	50085
0.3096	152	50237
0.3103	2	50239
0.3107	2	50241
0.3111	40	50281
0.3194	3	50284
0.325	80	50364
0.3269	1	50365
0.3333	29695	80060
0.3334	213	80273
0.3338	5	80278
0.3409	1	80279
0.3429	122	80401
0.3445	3	80404
0.35	2	80406
0.3512	1	80407
0.3572	5	80412
0.3576	1	80413
0.3651	1	80414
0.3667	363	80777
0.3678	1	80778
0.3767	2	80780
0.379	3	80783
0.3917	2	80785
0.3929	8	80793
0.3936	2	80795
0.3969	2	80798
0.4005	1	81535
0.402	5	81540
0.4028	1	81541
0.4047	4	81545
0.4083	1	81546
0.4102	6	81552
0.4108	8	81560
0.4166	17	81577
0.4167	3	81580
0.4207	1	81581
0.4222	1	81582
0.4242	24	81606
0.425	1	81607
0.4287	3	81610
0.4333	40	81650
0.4334	4	81654
0.4338	2	81656
0.4346	5	81661
0.4361	1	81662
0.4444	65	81727
0.4445	5	81732
0.45	6	81738
0.4525	11	81749
0.454	2	81751
0.4576	2	81753
0.4583	105	81858
0.4584	8	81866
0.4667	3	81869
0.4679	7	81876
0.4714	2	81878
0.4722	1	81879
0.4762	158	82037
0.4763	16	82053
0.4778	6	82059
0.4858	8	82067
0.4901	1	82068
0.4909	1	82069
0.4917	17	82086
0.5	27259	109345
0.5001	25	109370
0.5011	1	109371
0.5075	1	109372

...Continued

Fraction of Papers	No. of Authors	Cum. Freq.
z	f(z)	F(z)
0.5111	3	109388
0.5167	1	109389
0.5222	1	109390
0.5242	1	109391
0.525	15	109406
0.5255	1	109407
0.5277	2	109409
0.5333	896	110305
0.5334	41	110346
0.5429	21	110367
0.5444	1	110368
0.5448	1	110369
0.5525	3	110372
0.5531	1	110373
0.5537	2	110375
0.5555	3	110378
0.5596	1	110379
0.5603	1	110380
0.5636	1	110381
0.5667	44	110425
0.5671	1	110426
0.5694	3	110429
0.5714	1	110430
0.5716	1	110431
0.5762	1	110432
0.5775	3	110435
0.5833	2	110437
0.5858	2	110439
0.5873	6	110445
0.5874	1	110446
0.5909	1	110447
0.5969	1	110448
0.6	73	110521
0.6012	1	110522
0.6013	1	110523
0.602	1	110524
0.6102	1	110525
0.6108	1	110526
0.6111	6	110532
0.6191	12	110544
0.6192	1	110545
0.6198	1	110546
0.6222	1	110547
0.6242	4	110551
0.6243	1	110552
0.625	13	110565
0.6251	1	110566
0.6262	1	110567
0.63	1	110568
0.6334	2	110574
0.6338	1	110575
0.6346	1	110576
0.6429	14	110590
0.643	5	110595
0.6444	7	110602
0.6445	1	110603
0.6449	1	110604
0.65	1	110605
0.6525	2	110607
0.6531	1	110608
0.6583	10	110618
0.6648	1	110619
0.6666	3057	113676
0.6667	25	113701
0.6668	2	113703
0.6679	4	113707
0.6762	22	113729
0.6763	6	113735
0.6778	4	113739
0.6917	3	113742
0.6922	1	113743
0.6944	1	113744
0.7	62	113806
0.7001	2	113808
0.7025	1	113809
0.7096	1	113810
0.7111	1	113811
0.7333	108	113919
0.7334	4	113923
0.7353	2	113925
0.7401	1	113926
0.7429	5	113931
0.7499	4	113935
0.75	1713	115648
0.7525	3	115651
0.7575	2	115653
0.762	2	115655
0.7666	4	115659
0.7667	4	115663
0.7679	3	115666
0.7714	1	115667
0.7762	1	115668
0.7777	10	115678
0.7778	1	115679
0.7833	1	115680
0.7858	3	115683
0.7916	11	115694
0.8	17	115711
0.8095	40	115753
0.8096	3	115756
0.8131	1	115757
0.8192	1	115758
0.8194	1	115759
0.825	2	115761
0.83	1	115762
0.8333	64	115826
0.8339	1	115827
0.8346	3	115830
0.8353	1	115831
0.8429	5	115836
0.843	1	115837
0.8444	3	115840
0.8525	1	115841
0.8583	5	115846
0.8666	183	116029
0.8667	7	116036
0.8762	3	116039
0.8825	1	116040
0.8858	1	116041
0.887	3	116044
0.8928	2	116046
0.8936	1	116047
0.9	13	116060
0.9001	1	116061
0.9095	1	116062
0.9166	2	116064
0.9198	1	116065
0.9262	1	116066
0.9333	10	116076
0.9345	1	116077
0.9361	1	116078
0.9435	1	116079
0.9445	1	116080
0.9524	2	116082
0.9583	3	116085
0.9666	3	116088
0.9667	1	116089
0.9762	3	116092
0.9777	2	116094
0.9916	1	116095
0.9917	1	116096
0.9999	581	116677
1	9318	125995
1.0095	2	125997
1.0111	1	125998

...Continued

Fraction of Papers	No. of Authors	Cum. Freq.
z	f(z)	F(z)
1.0263	1	126000
1.0333	14	126014
1.0334	1	126015
1.0429	1	126016
1.043	1	126017
1.0493	1	126018
1.0583	1	126019
1.0584	1	126020
1.0666	36	126056
1.0667	1	126057
1.0714	1	126058
1.0769	1	126059
1.0873	1	126060
1.0909	6	126066
1.0953	1	126067
1.0999	1	126068
1.1	8	126076
1.1012	1	126077
1.1028	1	126078
1.111	1	126079
1.1111	10	126089
1.1191	1	126090
1.1192	1	126091
1.1242	1	126092
1.1249	2	126094
1.125	9	126103
1.1333	4	126107
1.1361	1	126108
1.1428	7	126115
1.1429	22	126137
1.1525	1	126138
1.1666	16	126154
1.1667	36	126190
1.1762	1	126191
1.1777	2	126193
1.188	1	126194
1.1916	3	126197
1.1999	53	126250
1.2	82	126332
1.2096	1	126333
1.2221	1	126334
1.225	1	126335
1.2251	1	126336
1.2333	5	126341
1.2382	1	126342
1.2435	1	126343
1.2499	1	126344
1.25	344	126688
1.2539	1	126689
1.2611	1	126691
1.2666	4	126695
1.2667	3	126698
1.2713	1	126699
1.2762	1	126700
1.2909	1	126701
1.2917	1	126702
1.3	1	126703
1.3095	1	126704
1.3096	2	126706
1.3194	1	126707
1.3249	2	126709
1.325	1	126710
1.3332	164	126874
1.3333	428	127302
1.3334	2	127304
1.3372	1	127305
1.3428	2	127307
1.3429	4	127311
1.3666	2	127313
1.3667	6	127319
1.3858	1	127320
1.3916	1	127321
1.3999	12	127333
1.4	9	127342
1.4012	1	127343
1.4019	1	127344
1.4077	1	127345
1.4095	1	127346
1.4123	1	127347
1.4241	1	127348
1.4242	1	127349
1.4334	2	127351
1.4443	1	127352
1.4444	1	127353
1.4582	1	127354
1.4583	4	127358
1.4666	3	127361
1.4667	1	127362
1.4761	5	127367
1.4762	3	127370
1.4778	1	127371
1.4857	2	127373
1.4999	8	127381
1.5	677	128058
1.5095	1	128059
1.5249	1	128060
1.5332	17	128077
1.5333	18	128095
1.5666	2	128098
1.5715	1	128099
1.5789	1	128100
1.5999	4	128104
1.6	3	128107
1.6095	1	128108
1.611	1	128109
1.6111	1	128110
1.6167	1	128111
1.625	1	128112
1.6333	2	128114
1.6665	62	128176
1.6666	75	128251
1.6667	1	128252
1.6762	1	128253
1.6999	3	128256
1.7	1	128257
1.7332	7	128264
1.7333	5	128269
1.738	1	128270
1.75	113	128383
1.7666	1	128384
1.7776	1	128385
1.7777	1	128386
1.7789	1	128387
1.8094	1	128388
1.8095	3	128391
1.8249	1	128392
1.8333	2	128394
1.8428	1	128395
1.8582	1	128396
1.8665	16	128412
1.8666	10	128422
1.9	1	128423
1.9332	2	128425
1.9691	1	128426
1.9762	1	128427
1.9776	1	128428
1.9998	20	128448
1.9999	27	128475
2	604	129079
2.0095	1	129080
2.0332	5	129085
2.0334	1	129086
2.0665	4	129090
2.0666	2	129092
2.0667	1	129093

...continued

Fraction of Papers	No. of Authors	Cum. Freq.
z	f(z)	F(z)
2.125	1	129095
2.1333	1	129096
2.1428	1	129097
2.1429	1	129098
2.1582	1	129099
2.1665	4	129103
2.1666	2	129105
2.1667	3	129108
2.1679	1	129109
2.1777	1	129110
2.1916	1	129111
2.1998	7	129118
2.1999	3	129121
2.2	5	129126
2.25	49	129175
2.2679	1	129176
2.2917	1	129177
2.311	1	129178
2.3111	1	129179
2.3331	10	129189
2.3332	9	129198
2.3333	42	129240
2.3998	3	129243
2.3999	2	129245
2.4	2	129247
2.4242	1	129248
2.4581	1	129249
2.4678	1	129250
2.4762	1	129251
2.4999	1	129252
2.5	85	129337
2.5331	1	129338
2.5332	2	129340
2.6	1	129341
2.6664	4	129345
2.6665	2	129347
2.6666	10	129357
2.676	1	129358
2.7331	1	129359
2.7333	1	129360
2.75	25	129385
2.7667	2	129387
2.8332	1	129388
2.8664	2	129390
2.8666	1	129391
2.9094	1	129392
2.9666	1	129393
2.9667	2	129395
2.976	1	129396
2.9998	1	129398
2.9999	2	129400
3	105	129505
3.0343	1	129506
3.0666	1	129507
3.0833	1	129508
3.0909	1	129509
3.1012	1	129510
3.1429	1	129511
3.1997	4	129515
3.1998	1	129516
3.2	2	129518
3.25	9	129527
3.2664	1	129528
3.333	1	129529
3.3331	1	129530
3.3332	1	129531
3.3333	7	129538
3.4583	1	129539
3.5	21	129560
3.533	1	129561
3.5331	1	129562
3.5333	1	129563
3.5998	1	129564
3.6663	1	129565
3.6666	3	129568
3.733	1	129569
3.75	7	129576
3.8663	1	129577
3.8759	1	129578
3.9999	3	129581
4	30	129611
4.0094	1	129612
4.25	6	129618
4.3333	2	129620
4.3667	1	129621
4.5	9	129630
4.5331	1	129631
4.5332	1	129632
4.6666	1	129633
5	15	129648
5.25	3	129651
5.2759	1	129652
5.5	4	129656
5.75	2	129658
6	5	129663
6.5	1	129664
6.5327	1	129665
7	6	129671
8	4	129676
8.3333	1	129677
9	4	129681
10	2	129683
11	1	129684
12	2	129686
13	2	129688
24	1	129689
	129689	
Mean:	0.430125	
Variance:	0.133416	
St. Dev.:	0.365262	

Table 3. Values of f(q) for case i = 2

q	Obs. Data	Nor. Obs. data	Lotka-Lotka	Nor. Lotka-Lotka	Poisson-Lotka	Nor. Poisson - Lotka	Neg. Bin.-Poisson	Nor. NB-Poisson
1/2	0.891733	0.8952816	0.477461	0.6021017	0.573485	0.7020321	0.813218	0.832419
1	0.085127	0.0854655	0.224192	0.2827175	0.16967	0.207702	0.129925	0.132993
3/4	0.01307	0.0131217	0.083952	0.105868	0.072944	0.0892939	0.033462	0.034252
2	0.006107	0.0061312	0.007385	0.0093129	0.000794	0.0009724	0.000328	0.000336
Total	0.996037	1	0.792991	1.0000001	0.816893	1.0000004	0.976933	1

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

q	Obs. Data	Nor. Obs. data	Lotka-Lotka	Nor. Lotka-Lotka	Poisson-Lotka	Nor. Poisson - Lotka	Neg. Bin.-Poisson	Nor. NB-Poisson
1/3	0.629043	0.632016	0.412241	0.483389	0.447461	0.508579	0.634513	0.638378
1/2	0.222509	0.223561	0.06522	0.076476	0.126024	0.143237	0.178706	0.179794
2/3	0.040181	0.040371	0.069889	0.081951	0.082341	0.093588	0.049386	0.049687
5/6	0.003208	0.003223	0.022114	0.025931	0.046382	0.052717	0.027819	0.027988
1	0.077069	0.077433	0.153253	0.179703	0.067885	0.077157	0.06399	0.064379
7/6	0.00485	0.004873	0.009998	0.011723	0.02276	0.025869	0.009522	0.00958
4/3	0.005081	0.005105	0.053845	0.063138	0.03023	0.034359	0.013415	0.013497
3/2	0.005806	0.005834	0.012165	0.014265	0.016734	0.01902	0.005925	0.005961
5/3	0.002182	0.002192	0.024689	0.02895	0.016778	0.01907	0.005045	0.005075
11/6	0.000308	0.00031	0.009213	0.010803	0.011897	0.013522	0.003085	0.003104
2	0.005058	0.005082	0.020188	0.023672	0.011335	0.012883	0.002541	0.002556
Total	0.995296	1	0.852815	1	0.879827	1	0.993945	1

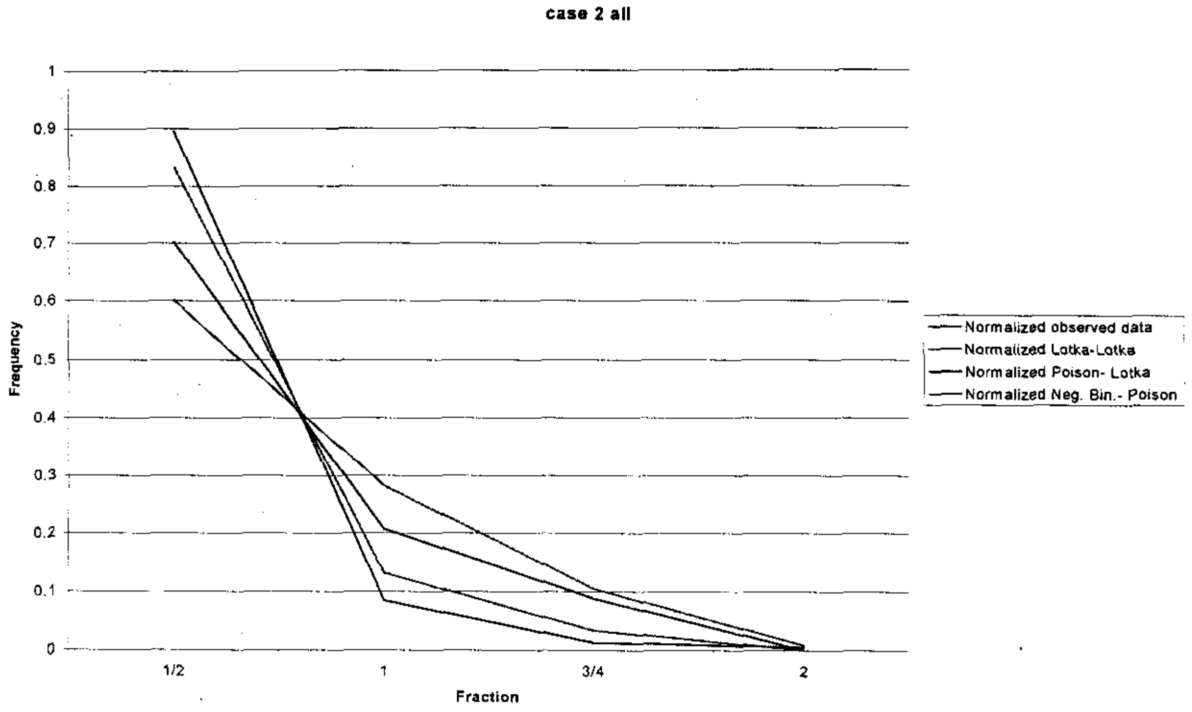
Table 5. Values of f(q) for case i = 4

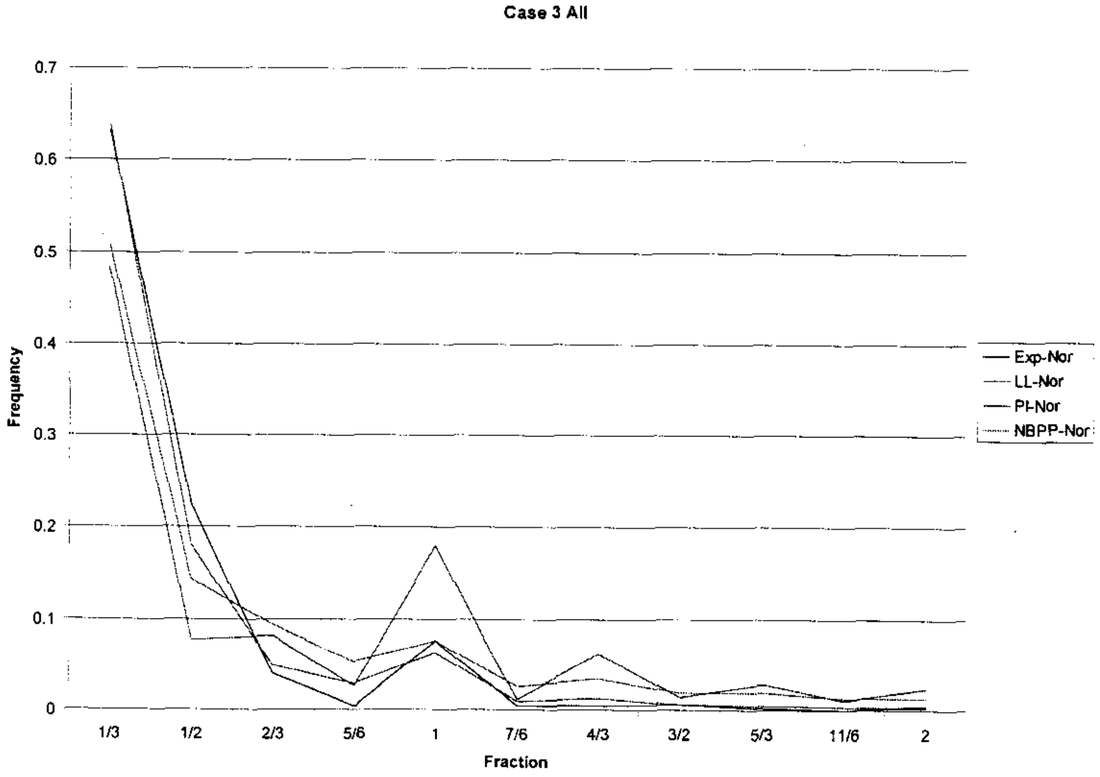
q	Obs. Data	Nor. Obs. data	Lotka-Lotka	Nor. Lotka-Lotka	Poisson-Lotka	Nor. Poisson - Lotka	Neg. Bin.-Poisson	Nor. NB-Poisson
1/4	0.385931	0.38773	0.368761	0.4301577	0.274408	0.308167	0.389119	0.391061
1/3	0.236928	0.23804	0.04348	0.0507191	0.173053	0.194342	0.245394	0.246619
1/2	0.227992	0.22906	0.121144	0.1413135	0.156991	0.176305	0.197279	0.198264
7/12	0.001689	0.0017	0.013188	0.0153835	0.039058	0.043863	0.023426	0.023543
2/3	0.025014	0.02513	0.000777	0.0009069	0.012316	0.013831	0.007387	0.007424
3/4	0.014535	0.0146	0.034859	0.0406629	0.034657	0.03892	0.019659	0.019757
5/6	0.002637	0.00265	0.007666	0.008942	0.029692	0.033345	0.015676	0.015754
11/12	0.000378	0.00038	0.000629	0.0007335	0.007413	0.008324	0.003101	0.003117
1	0.076568	0.07693	0.145334	0.1695309	0.051085	0.057369	0.056745	0.057028
13/12	0.000679	0.00068	0.004313	0.005031	0.014776	0.016593	0.005636	0.005664
7/6	0.001766	0.00177	0.00054	0.0006303	0.007169	0.00805	0.002483	0.002495
5/4	0.002853	0.00287	0.046649	0.0544157	0.016635	0.018682	0.007646	0.007684
4/3	0.004758	0.00478	0.008132	0.0094865	0.02121	0.023819	0.007721	0.00776
17/12	0.00027	0.00027	0.000429	0.000501	0.00527	0.005918	0.001301	0.001307
3/2	0.005714	0.00574	0.023194	0.0270554	0.009779	0.010982	0.003478	0.003495
19/12	0.000131	0.00013	0.005431	0.0063351	0.009007	0.010115	0.002546	0.002559
5/3	0.001118	0.00112	0.000559	0.0006517	0.005119	0.005748	0.001169	0.001175
7/4	0.001002	0.00101	0.014984	0.0174793	0.006718	0.007545	0.001785	0.001794
11/6	0.00027	0.00027	0.001834	0.0021398	0.006214	0.006979	0.001424	0.001431
23/12	2.31E-05	2.3E-05	0.000563	0.0006572	0.004138	0.004647	0.000687	0.00069
2	0.005097	0.00512	0.014803	0.0172676	0.005747	0.006454	0.001373	0.00138
Total	0.99535	1	0.85727	1.0000006	0.890454	1	0.995034	1

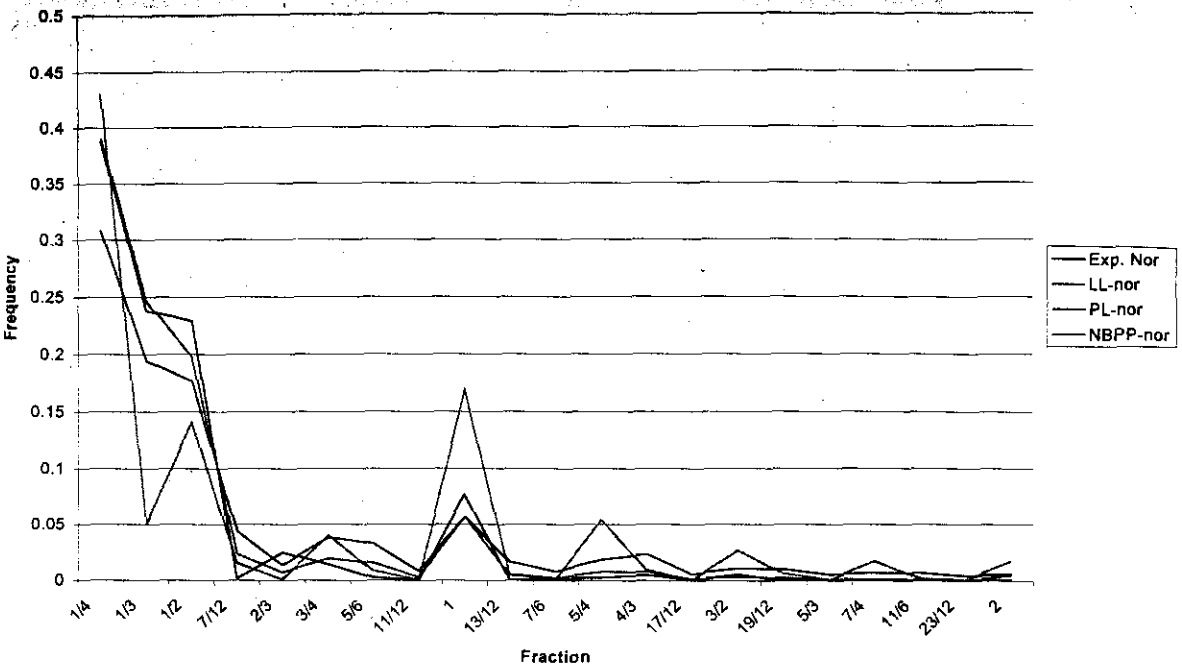
Table 6. Values of f(q) for case i = 5								
q	Obs. Data	Nor. Obs. data	Lotka-Lotka	Nor. Lotka-Lotka	Poisson-Lotka	Nor. Poisson - Lotka	Neg. Bin.-Poisson	Nor. NB-Poisson
12/60	0.210272	0.211268	0.336166	0.387328	0.133595	0.15027	0.189433	0.190367
15/60	0.17249	0.173306	0.032611	0.037575	0.140825	0.15841	0.199685	0.200669
20/60	0.234561	0.235671	0.043482	0.050099	0.17306	0.19467	0.245394	0.246603
24/60	0.01145	0.011505	0.046472	0.053545	0.00734	0.00826	0.004402	0.004424
27/60	0.001588	0.001596	0.009017	0.010389	0.015473	0.01741	0.00928	0.009326
30/60	0.212971	0.213979	0.06566	0.075653	0.134185	0.15094	0.183597	0.184502
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
35/60	0.000702	0.000705	0.001166	0.001344	0.020045	0.02255	0.012022	0.012081
36/60	0.000671	0.000674	0.011421	0.01316	0.000717	0.00081	0.0003	0.000301
39/60	0.000648	0.000651	0.003324	0.00383	0.002267	0.00255	0.000948	0.000953
40/60	0.023896	0.024009	0.000778	0.000896	0.012316	0.01385	0.007387	0.007423
42/60	0.000786	0.00079	0.018356	0.021149	0.016237	0.01826	0.009305	0.009351
44/60	0.000879	0.000883	0.004432	0.005106	0.002786	0.00313	0.001165	0.001171
45/60	0.013324	0.013387	0.00176	0.002028	0.015437	0.01736	0.009106	0.009151
47/60	0.000224	0.000225	0.00086	0.000991	0.005873	0.00661	0.002457	0.002469
48/60	0.000254	0.000256	0.003553	0.004093	8.86E-05	1E-04	2.49E-05	2.5E-05
50/60	0.000902	0.000906	0.002374	0.002736	0.021034	0.02366	0.012054	0.012113
51/50	8.48E-05	8.52E-05	0.001379	0.001588	0.000374	0.00042	0.000105	0.000106
52/60	0.001504	0.001511	0.000573	0.00066	0.003609	0.00406	0.00151	0.001517
54/60	0.000193	0.000194	0.006848	0.007891	0.00262	0.00295	0.001015	0.00102
55/60	3.08E-05	3.1E-05	5.56E-05	6.41E-05	0.003804	0.00428	0.001591	0.001599
56/60	9.25E-05	9.3E-05	0.001838	0.002118	0.000459	0.00052	0.000129	0.00013
57/60	3.86E-05	3.87E-05	0.001303	0.001501	0.004692	0.00528	0.001906	0.001915
59/60	9.25E-05	9.3E-05	0.000535	0.000616	0.001452	0.00163	0.000408	0.00041
60/60	0.076344	0.076706	0.13354	0.153864	0.044884	0.05049	0.054348	0.054616
62/60	0.000154	0.000155	0.001772	0.002041	0.006787	0.00763	0.002629	0.002642
63/60	2.31E-05	2.32E-05	0.00061	0.000703	6.57E-05	7.4E-05	1.19E-05	1.2E-05
64/60	0.000308	0.00031	0.000357	0.000411	0.000892	0.001	0.000251	0.000252
65/60	7.71E-06	7.75E-06	0.000169	0.000194	0.006078	0.00684	0.002469	0.002481
66/60	0.000154	0.000155	0.002875	0.003313	0.000473	0.00053	0.000119	0.00012
67/60	0.000123	0.000124	6.92E-05	7.97E-05	0.001881	0.00212	0.000529	0.000532
68/60	0.0001	0.000101	0.000813	0.000937	8.07E-05	9.1E-05	1.46E-05	1.47E-05
69/60	0.000239	0.00024	0.000814	0.000938	0.001203	0.00135	0.000324	0.000325
70/60	0.000401	0.000403	0.000115	0.000132	0.004396	0.00494	0.001703	0.001711
71/60	2.31E-05	2.32E-05	0.000316	0.000364	0.00034	0.00038	6.16E-05	6.19E-05
72/60	0.001072	0.001077	0.037948	0.043723	0.00766	0.00862	0.003613	0.003631
74/60	6.94E-05	6.97E-05	0.001116	0.001286	0.001838	0.00207	0.000463	0.000465
75/60	0.002683	0.002696	0.00391	0.004506	0.007236	0.00814	0.003578	0.003595
76/60	7.71E-05	7.75E-05	0.00021	0.000242	0.000209	0.00024	3.79E-05	3.81E-05
77/60	7.71E-06	7.75E-06	0.000211	0.000243	0.003118	0.00351	0.000839	0.000843
78/60	2.31E-05	2.32E-05	0.001288	0.001484	9.05E-05	0.0001	1.42E-05	1.43E-05
79/60	4.63E-05	4.65E-05	6.12E-05	7.05E-05	0.000661	0.00074	0.00012	0.00012
80/60	0.004588	0.00461	0.005217	0.006011	0.008936	0.01005	0.004401	0.004423
81/60	5.4E-05	5.42E-05	0.000482	0.000555	0.000292	0.00033	4.99E-05	5.02E-05
82/60	6.17E-05	6.2E-05	0.000144	0.000166	0.00238	0.00268	0.0006	0.000603
83/60	0	0	0.000182	0.000209	7.79E-05	8.8E-05	8.79E-06	8.83E-06

Continued...

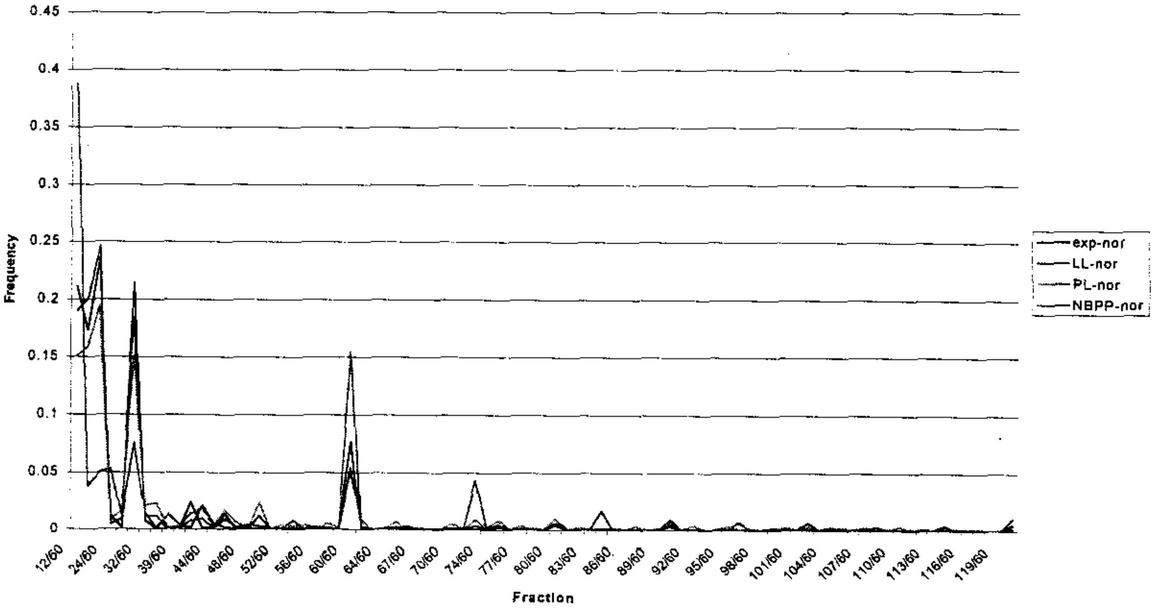
q	Obs. Data	Nor. Obs. data	Lotka-Lotka	Nor. Lotka-Lotka	Poisson-Lotka	Nor. Poisson - Lotka	Neg. Bin.-Poisson	Nor. NB-Poisson
84/60	0.000193	0.000194	0.014391	0.016581	0.001726	0.00194	0.000489	0.000491
85/60	3.86E-05	3.87E-05	1.36E-05	1.57E-05	0.002019	0.00227	0.000543	0.000546
86/60.	1.54E-05	1.55E-05	0.000666	0.000768	0.000469	0.00053	7.37E-05	7.4E-05
87/60	1.54E-05	1.55E-05	0.002878	0.003316	0.003029	0.00341	0.000924	0.000929
88/60	6.94E-05	6.97E-05	0.000121	0.00014	4.79E-05	5.4E-05	5.4E-06	5.43E-06
89/60	6.94E-05	6.97E-05	0.000187	0.000216	0.001136	0.00128	0.000194	0.000195
90/60	0.005297	0.005322	0.007821	0.009011	0.006432	0.00723	0.003071	0.003086
91/60	1.54E-05	1.55E-05	4.7E-05	5.42E-05	0.000202	0.00023	2.28E-05	2.29E-05
92/60	0.00027	0.000271	0.003844	0.004429	0.003946	0.00444	0.001171	0.001177
93/60	7.71E-06	7.75E-06	0.000279	0.000321	6.93E-05	7.8E-05	7.26E-06	7.3E-06
94/60	2.31E-05	2.32E-05	0.000129	0.000149	0.000911	0.00102	0.000143	0.000144
95/60	7.71E-06	7.75E-06	0.000458	0.000528	0.003384	0.00381	0.001098	0.001103
96/60	5.4E-05	5.42E-05	0.006145	0.00708	0.000418	0.00047	6.95E-05	6.98E-05
97/60	3.08E-05	3.1E-05	2.42E-05	2.79E-05	0.001472	0.00166	0.000251	0.000252
98/60	2.31E-05	2.32E-05	0.000389	0.000448	0.000116	0.00013	1.1E-05	1.11E-05
99/60	0	0	0.001818	0.002095	0.001036	0.00117	0.000189	0.00019
100/60	0.001064	0.001069	0.000306	0.000353	0.002497	0.00281	0.000747	0.000751
101/60	7.71E-06	7.75E-06	0.000144	0.000166	0.000359	0.0004	3.76E-05	3.78E-05
102/60	3.08E-05	3.1E-05	0.005732	0.006604	0.002877	0.00324	0.000805	0.000809
103/60	0	0	3.34E-05	3.85E-05	5.7E-05	6.4E-05	3.91E-06	3.93E-06
104/60	9.25E-05	9.3E-05	0.002432	0.002802	0.001383	0.00156	0.000243	0.000244
105/60	0.000879	0.000883	0.000676	0.000779	0.002319	0.00261	0.000714	0.000717
106/60	7.71E-06	7.75E-06	0.000101	0.000116	0.000301	0.00034	2.86E-05	2.88E-05
107/60	2.31E-05	2.32E-05	0.000511	0.000589	0.002179	0.00245	0.000433	0.000435
108/60	0	0	0.002801	0.003227	0.000103	0.00012	9.87E-06	9.92E-06
109/60	3.86E-05	3.87E-05	2.8E-05	3.23E-05	0.000698	0.00079	7.31E-05	7.35E-05
110/60	1.54E-05	1.55E-05	0.000927	0.001068	0.002986	0.00336	0.000924	0.000929
111/60	7.71E-06	7.75E-06	0.001095	0.001261	0.000325	0.00037	3.45E-05	3.47E-05
112/60	0.0002	0.000201	0.000342	0.000394	0.001731	0.00195	0.000308	0.000309
113/60	0	0	0.000103	0.000119	0.000105	0.00012	6.59E-06	6.62E-06
114/60	7.71E-06	7.75E-06	0.003609	0.004158	0.001069	0.0012	0.00017	0.000171
115/60	0	0	5.19E-05	5.98E-05	0.001358	0.00153	0.000273	0.000275
116/60	1.54E-05	1.55E-05	0.001467	0.00169	0.000443	0.0005	4.49E-05	4.51E-05
117/60	0	0	0.000743	0.000856	0.001583	0.00178	0.000285	0.000286
118/60	7.71E-06	7.75E-06	7.23E-05	8.34E-05	9.15E-05	0.0001	5.15E-06	5.17E-06
119/60	7.71E-06	7.75E-06	0.000444	0.000511	0.000975	0.0011	0.000114	0.000115
120/60	0.00502	0.005043	0.008879	0.01023	0.002271	0.00255	0.00076	0.000764
Total	0.995289	1	0.86791	1	0.889005	1	0.995096	1


case 2 all
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Frequency
1/2
1
3/4
2
Fraction
Normalized observed data
Normalized Lotka-Lotka
Normalized Poison- Lotka
Normalized Neg. Bin.- Poison


Case 3 All
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Frequency
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
Pl-Nor
NBPP-Nor



Case 4 all



Case 5 All