Department of Library and Information Science

Bharathidasan University

Tiruchirappalli-620024

Name of the Programme: M.Lib.I.Sc

Course - 4.5: Elective – V INFORMETRICS AND SCIENTOMETRICS

Course Code:P21MLS20A

Unit-II: Theory and Laws - Zipf's law, Lotka's Law, Bradford's Law. Price Theory

Presented

By

Course Teacher: DR.C.RANGANATHAN

Professor

e-mail: cranganathan72@gmail.com

Bibliometric Laws & Methods

- Lotka's law (1926)
  - Productivity of authors (researchers); how many researchers have written 1, 2, 3… articles?
  - Frequency distribution (probability)
- Bradford's law (1934)
  - Scatter of the articles of a given subject over journals
  - *equal-production groups*
- Zipf's law (1949)
  - The frequency of words in text
  - Rank/Frequency distribution
- Impact factor
- Citation structures
- Co-citation structures

Alfred J. Lotka 1926

- Statistics—the frequency distribution of scientific productivity.
- Lotka's law of authorship describes the publication frequencies for authors within a given domain
- " . . . the number (of authors) making ***n*** contributions is about **1/*n*²** of those making one; … and the proportion of all contributors, that make a single contribution, is about 60 percent"
- $f(n) = C / n^2$ - where $f(n)$ denotes the number of authors within a domain who, in a given period, have published ***n*** articles
- ***C*** is a constant and denotes the number of researchers that have produced 1 article
- The number of researchers that have published ***n*** articles is proportional with $1 / n^2$ – that is …

Cont…,

- If 100 researchers have published one article each within a given domain then Lotka’s law predicts that:
- **$100 / 2^2 = 25$ researchers**, will have published 2 articles, and
- $100/3^2 = 11$ **researchers**, will have published 3 articles, etc.
- (This may be tested empirically by Dialog’s RANK command…)
- If Lotka’s law is used on a suitably large corpus and over a fairly long period of time, the result will generally be quite precise, but not statistically exact

Lotka's law: $x^n \bullet y = C$

The total number of authors *y* in a given subject, each producing *x* publications, is inversely proportional to some exponential function *n* of *x*.

- Where:
  - *x* = number of publications
  - *y* = no. of authors credited with *x* publications
  - *n* = constant (equals 2 for scientific subjects)
  - *C* = constant
- inverse square law of scientific productivity

Year wise Distribution of Research Output

Sl. No.	Publication Year	Records	%	TLCS	TGCS	ACPP	h-Index	CR	ACRPA	NA	AAPA
1	1991	49	0.4	20	1664	33.95	20	1310	26.73	142	2.90
2	1992	53	0.5	85	2517	47.49(3)	24	1921	36.25	157	2.96
3	1993	59	0.5	56	3319	56.25(2)	25	2750	46.61	206	3.49
4	1994	51	0.4	44	1732	33.96	19	1856	36.39	151	2.96
5	1995	75	0.7	79	2468	32.90	28	2437	32.49	252	3.36
6	1996	81	0.7	72	2806	34.64	27	3037	37.49	240	2.96
7	1997	85	0.7	74	3344	39.34	29	3448	40.56	267	3.14
8	1998	109	0.9	103	4514	41.41	33	4279	39.26	320	2.94
9	1999	135	1.2	467	4016	29.74	33	5015	37.15	385	2.85
10	2000	196	1.7	656	6057	30.90	44	6068	30.96	593	3.03
11	2001	221	1.9	760	10107	45.73	56	8214	37.17	754	3.41
12	2002	261	2.3	1389	14992	57.44(1)	58	8327	31.90	803	3.08
13	2003	328	2.9	666	11860	36.15	60	11529	35.15	1376	4.20
14	2004	363	3.2	877	13282	36.58	60	13617	37.51	1291	3.56
15	2005	473	4.1	1437	16984	35.90	63	18466	39.04	1712	3.62
16	2006	559	4.9	1522	19087	34.14	72(2)	22125	39.58	2054	3.67
17	2007	636	5.5	1799(1)	21193 (2)	33.32	75(1)	27201	42.77	2353	3.70
18	2008	693	6.0	1745 (2)	19109	27.57	66	27770	40.07	2698	3.89
19	2009	842	7.3	1430	21102 (3)	25.06	69(3)	39824	47.30,	3282	3.90
20	2010	879	7.6	1452(3)	22204 (1)	25.26	68	44558	50.69,	3741	4.26
21	2011	1104	9.6	1154	16968	15.36	55	57443	52.03,	4611	4.18
22	2012	1251(3)	10.9	1206	16806	13.43	51	64512(3)	51.57	5278(3)	4.22
23	2013	1424(2)	12.4	581	10178	7.14	35	74202(2)	52.11	6107(2)	4.29
24	2014	1576(1)	13.7	133	5819	3.69	24	84548(1)	53.65	6994(1)	4.44
Total		11503	100	17807	252128	21.91	1094	534457	824.41	45767	85.01

Author wise Distributions of publications

Year	Single	Two	Three	Four	Five	Six	Seven	Eight	Nine	Ten and Above	Total
1991	12	8	12	10	4	3	-	-	-	-	49
1992	8	15	14	10	2	1	3	-	-	-	53
1993	6	13	13	16	6	2	1	-	-	2	59
1994	12	11	11	11	3	1	-	1	-	1	51
1995	14	18	14	14	4	4	4	-	-	3	75
1996	16	30	8	13	6	4	2	-	1	1	81
1997	19	14	23	14	7	2	3	1	1	1	85
1998	20	27	24	21	13	3	1	-	-	-	109
1999	49	20	19	25	11	5	2	1	1	2	135
2000	57	38	32	33	17	8	4	1	2	4	196
2001	52	37	40	38	22	10	13	2	2	5	221
2002	65	48	60	37	25	12	6	4	2	2	261
2003	54	58	72	57	43	23	8	4	3	6	328
2004	64	64	71	67	43	20	14	11	3	6	363
2005	68	92	98	71	62	44	22	7	4	5	473
2006	78	112	113	86	72	43	22	18	6	9	559
2007	74	130	128	115	72	54	39	13	4	7	636
2008	88	123	135	112	98	57	33	19	12	16	693
2009	101(2)	140	167	155	107	71	46	28	15	12	842
2010	61	150	172	168	120	83	55	26	20	24	879
2011	104(1)	170	216	204	155	101	71	37	20	26	1104
2012	91(3)	212(3)	223(3)	255(2)	174	140	57	44	21	34	1251
2013	85	246(1)	283(2)	245(3)	204	162	82	52	22	43	1424
2014	84	231(2)	288(1)	294(1)	248	190	103	58	37	43	1576
Total	1282	2007(3)	2236(1)	2071(2)	1518	1043	591	327	176	252	11503
Mean	53.42	83.63(3)	93.17(1)	86.29(2)	63.25	43.46	26.86	18.17	9.78	12.00	479.29
S.D.	32.26	75.39	89.53	88.14	71.21	55.18	30.07	18.72	10.41	13.74	472.74
Co-efficient variation	0.60	0.90	0.96	1.02	1.13	1.27	1.12	1.03	1.06	1.14	0.99

Ranking of Prolific Authors based on Publications

Sl. No	Author	Records	Percentage	TLCS	TGCS	TLCR	h-index
1	**Mallakpour S**	**64 (1)**	0.6	259	956	**285 (2)**	19
2	**Anastas PT**	**60 (2)**	0.5	**1115 (1)**	**2640 (3)**	126	20
3	**Clark JH**	**49 (3)**	0.4	**589 (2)**	**3043 (1)**	154	**26(1)**
4	**Warner JC**	44	0.4	74	208	20	8
5	**Varma RS**	39	0.3	511	**2804 (2)**	**294 (1)**	**26(1)**
6	Tu SJ	35	0.3	80	663	69	13
7	Zhang Y	34	0.3	28	404	31	12
8	**Kirchhoff MM**	32	0.3	**568 (3)**	1322	18	7
9	**Mizuno N**	31	0.3	157	1718	148	**22(3)**
10	Wang L	31	0.3	24	803	25	12
11	Yadav GD	31	0.3	53	333	53	10
12	**Yamaguchi K**	31	0.3	172	1873	150	**23(2)**
13	Heravi MM	30	0.3	74	847	63	15
14	Hutchison JE	28	0.2	93	219	29	9
15	Poliakoff M	27	0.2	328	1156	55	14
16	Tundo P	27	0.2	260	1065	102	12
17	Han BX	26	0.2	52	637	59	14
18	Liu Y	24	0.2	8	382	48	12
19	Azizi N	23	0.2	61	323	65	10
20	Luque R	23	0.2	79	645	74	16
21	Singh S	23	0.2	33	271	47	8
22	Garcia H	22	0.2	93	1756	81	18
23	Kidwai M	22	0.2	33	140	32	8
24	Kumar S	22	0.2	42	250	37	9
25	Li CJ	22	0.2	201	1233	59	15

Cont..,

Sl. No	Author	Records	%	TLCS	TGCS	TLCR	h-index
26	Liang XZ	22	0.2	20	71	128	6
27	Namiesnik J	22	0.2	158	528	174	16
28	Wang Y	22	0.2	36	590	63	13
29	Brush EJ	21	0.2	0	0	0	0
30	He LN	21	0.2	114	930	168	17
31	Vaccaro L	21	0.2	131	589	166	16
32	Wang J	21	0.2	31	583	20	11
33	Dandia A	20	0.2	39	334	43	12
34	Hasaninejad A	20	0.2	42	274	38	10
35	Rogers RD	20	0.2	103	1462	37	10
36	Astruc D	19	0.2	29	592	37	14
37	Chen J	19	0.2	86	1304	32	10
38	Collins TJ	19	0.2	90	744	26	11
39	Corma A	19	0.2	70	1362	60	14
40	Lipshutz BH	19	0.2	83	399	65	13
41	Oelgemoller M	19	0.2	125	429	133	11
42	**Polshettiwar V**	19	0.2	358	**2073(1)**	173	17
43	Zhang L	19	0.2	3	189	21	8
44	Cannon AS	18	0.2	45	112	4	3
45	de la Guardia M	18	0.2	120	492	54	9
46	**Hayashi S**	18	0.2	313	**1902(2)**	109	16
47	Kumar R	18	0.2	54	653	42	9
48	Li Y	18	0.2	5	284	14	7
49	Li Z	18	0.2	49	255	56	8
50	Macquarrie DJ	18	0.2	103	938	40	14

Lotka's Law of Author Productivity- Chi- Square Model

No. of authors	Observed Number of authors with 'n' or (an) or (f)	Observed percentage of authors 100 x an/a1	Expected number of authors (an=an/n^2)or (p)	Expected percentage of authors	(F-P)^2/P
1	1282	100	1282	100	0
2	2007	156.55	320.5	25	8874.51
3	2236	174.41	142.44	11.11	30769.73
4	2071	161.54	80.12	6.25	49467.5
5	1518	118.40	51.28	4	41951.4
6	1043	81.35	35.61	2.77	28497.63
7	591	46.09	26.16	2.04	12194.21
8	327	25.50	20.03	1.56	4704.14
9	176	13.72	15.82	1.23	1620.96
10	94	7.33	12.82	1	514.05
11	60	4.68	10.59	0.82	230.37
12	27	2.10	8.90	0.69	36.78
13	21	1.63	7.58	0.59	23.72
14	12	0.93	6.54	0.51	4.55
15	14	1.09	5.69	0.44	12.09
16	2	0.15	5.00	0.39	1.80
17	5	0.39	4.43	0.34	0.07
18	3	0.23	3.95	0.30	0.23
19	3	0.23	3.55	0.27	0.08
20	1	0.07	3.20	0.25	1.51
21	3	0.23	2.90	0.22	0.00
23	1	0.07	2.42	0.18	0.83
25	1	0.07	2.05	0.16	0.53
26	1	0.07	1.89	0.14	0.42
27	1	0.07	1.75	0.13	0.32
36	1	0.07	0.98	0.07	0.00
48	1	0.07	0.55	0.04	0.35
243	1	0.07	0.02	0.00	44.08
				χ^2	178951.9

Lotka's inverse square law of scientific Author Productivity

No. of Contribution X	No of Contributors	Y	∑X= log x	∑Y= log y	∑X*∑Y	∑X*X
1	24535	24535	0	10.107	0	0
2	3774	7548	0.693	8.929	6.189	0.480
3	1242	3726	1.098	8.223	9.033	1.206
4	507	2028	1.386	7.614	10.556	1.921
5	297	1485	1.609	7.303	11.754	2.590
6	183	1098	1.791	7.001	12.544	3.210
7	110	770	1.945	6.646	12.933	3.786
8	86	688	2.079	6.533	13.586	4.324
9	62	558	2.197	6.324	13.896	4.827
10	38	380	2.302	5.940	13.677	5.301
11	27	297	2.397	5.693	13.652	5.749
12	22	264	2.484	5.575	13.855	6.174
13	13	169	2.564	5.129	13.157	6.578
14	16	224	2.639	5.411	14.281	6.964
15	11	165	2.708	5.105	13.827	7.333
16	10	160	2.772	5.075	14.071	7.687
17	15	255	2.833	5.541	15.699	8.027
18	12	216	2.890	5.375	15.536	8.354
19	8	152	2.944	5.023	14.792	8.669
20	3	60	2.995	4.094	12.265	8.974
21	4	84	3.044	4.430	13.489	9.269
22	7	154	3.091	5.036	15.569	9.554
23	3	69	3.135	4.234	13.276	9.831
24	1	24	3.178	3.178	10.100	10.100
26	1	26	3.258	3.258	10.615	10.615
27	2	54	3.295	3.988	13.147	10.862
28	1	28	3.332	3.332	11.103	11.103
30	1	30	3.401	3.401	11.568	11.568
31	4	124	3.433	4.820	16.552	11.792
32	1	32	3.465	3.465	12.011	12.011
34	1	34	3.526	3.526	12.435	12.435
35	1	35	3.555	3.555	12.640	12.640
39	1	39	3.663	3.663	13.421	13.421
44	1	44	3.784	3.784	14.320	14.320
49	1	49	3.891	3.891	15.146	15.146
60	1	60	4.094	4.094	16.763	16.763
64	1	64	4.158	4.158	17.296	17.296
799	31003	45728	101.629	192.454	474.754	310.88

Price's Square Root Law for Food Chemistry Research Productivity

No.of Contribution A	No.of Contributors B	% of 31003	A*B	% of A*B	% Cumulated value of A*B	Cumulated value of A*B
64	1	0.003	64	0.14	0.14	64
60	1	0.003	60	0.13	0.27	124
49	1	0.003	49	0.11	0.38	173
44(217)	1	0.003 (0.012)	44 (217)	0.10 (0.48)	0.48	217 (0.474)
39	1	0.003	39	0.09	0.57	256
35	1	0.003	35	0.08	0.65	291
34	1	0.003	34	0.07	0.72	325
32	1	0.003	32	0.07	0.79	357
31	4	0.012	124	0.27	1.06	481
30	1	0.003	30	0.07	1.13	511
28	1	0.003	28	0.06	1.19	539
27	2	0.006	54	0.12	1.31	593
26	1	0.003	26	0.06	1.37	619
24	1	0.003	24	0.05	1.42	643
23	3	0.010	69	0.15	1.57	712
22	7	0.022	154	0.34	1.91	866
21	4	0.013	84	0.18	2.09	950
20	3	0.010	60	0.13	2.22	1010
19	8	0.025	152	0.33	2.55	1162
18	12	0.039	216	0.47	3.02	1378
17	15	0.048	255	0.56	3.58	1633
16	10	0.032	160	0.35	3.93	1793
15	11	0.035	165	0.36	4.29	1958
14	16	0.052	224	0.49	4.78	2182
13	13	0.042	169	0.37	5.15	2351
12	22	0.071	264	0.58	5.73	2615
11	27	0.087	297	0.65	6.38	2912
10	38	0.122	380	0.83	7.21	3292
9	62	0.199	558	1.22	8.43	3850
8	86	0.277	688	1.5	9.93	4538
7	110	0.354	770	1.68	11.61	5308
6	183	0.590	1098	2.4	14.01	6406
5	297	0.957	1485	3.25	17.26	7891
4	507	1.635	2028	4.43	21.69	9919
3	1242	4.006	3726	8.15	29.84	13645
2	3774	12.173	7548	16.51	46.35	21193
1	24535	79.137	24535	53.65	100	45728
799	31003	100	45728	100		

Pareto Principle (80 X 20 Rule) Law for Green Chemistry Research Productivity

No.of Contribution A	No.of Contributors B	% of 31003	A*B	% of A*B	% Cumulated value of A*B	Cumulated value of A*B
64	1	0.003	64	0.14	0.14	64
60	1	0.003	60	0.13	0.27	124
49	1	0.003	49	0.11	0.38	173
44	1	0.003	44	0.10	0.48	217
39	1	0.003	39	0.09	0.57	256
35	1	0.003	35	0.08	0.65	291
34	1	0.003	34	0.07	0.72	325
32	1	0.003	32	0.07	0.79	357
31	4	0.012	124	0.27	1.06	481
30	1	0.003	30	0.07	1.13	511
28	1	0.003	28	0.06	1.19	539
27	2	0.006	54	0.12	1.31	593
26	1	0.003	26	0.06	1.37	619
24	1	0.003	24	0.05	1.42	643
23	3	0.010	69	0.15	1.57	712
22	7	0.022	154	0.34	1.91	866
21	4	0.013	84	0.18	2.09	950
20	3	0.010	60	0.13	2.22	1010
19	8	0.025	152	0.33	2.55	1162
18	12	0.039	216	0.47	3.02	1378
17	15	0.048	255	0.56	3.58	1633
16	10	0.032	160	0.35	3.93	1793
15	11	0.035	165	0.36	4.29	1958
14	16	0.052	224	0.49	4.78	2182
13	13	0.042	169	0.37	5.15	2351
12	22	0.071	264	0.58	5.73	2615
11	27	0.087	297	0.65	6.38	2912
10	38	0.122	380	0.83	7.21	3292
9	62	0.199	558	1.22	8.43	3850
8	86	0.277	688	1.5	9.93	4538
7	110	0.354	770	1.68	11.61	5308
6	183	0.590	1098	2.4	14.01	6406
5	297	0.957	1485	3.25	17.26	7891
4(9919)	**507**	**1.635**	**2028 (9919)**	**4.43**	**21.69**	**9919 (21.69)**
3	1242	4.006	3726	8.15	29.84	13645
2	3774	12.173	7548	16.51	46.35	21193
1	24535	79.137	24535 (35809)	53.65	100	45728 (78.31)
799	**31003**	**100**	**45728**	**100**		

Samuel Clement Bradford 1934, 1948

- Distribution of quantity vs yield of sources of information on specific subjects
  - he studied **journals** as sources, but applicable to other
  - what journals produce how many articles in a subject and how are they distributed? or
  - How are articles in a subject scattered across journals?
- Purpose: to develop a method for identification of the most productive journals in a subject & deal with what he called "documentary chaos"

First published in: *Engineering* (1934) 137:85-86, then in his book *Documentation,* (1948)

Bradford's law

- Bradford's law of scattering is often used as a guideline to determine the number of core journals within a given subject
  - Most of the articles on a given subject have a tendency to be bunched together in a small number of journals
  - The rest tends to scattered over a large number of journals

Bradford’s law (verbal)

- **Two different wordings: a verbal and a graphical**

“If scientific journals are arranged in order of decreasing productivity of articles on a given subject, they may be divided into a nucleus of periodicals more particularly devoted to the subject and several groups or zones containing the same number of articles as the nucleus, when the number of periodicals in the nucleus and succeeding zones will be as **$1 : n : n^2$**”

(S. C. Bradford, 1934)

Bradford's law

- Bradford's law is usually presented in equal-production groups form, with this special procedure:
  1. rank the journal-titles in order of their productivity (i.e. in order of the frequency with which they publish articles on the specified subject)
  2. divide the journal-titles into ***m*** groups or zones, where each zone produces **1/*m*** of the total number of articles
  3. then the proportion of journals in each zone will be definable by the progression $n^0:n^1:n^2$ ..., where n is the *Bradford multiplier* → **$1 : n : n^2$**

Bradford's law

- e.g. $m = 3$ zones:
  - top 8 journals produce 110 articles
  - next 29 journals produce 133 articles
  - next 127 journals produce 152 articles
  - number of journals in each zone follows the progression 8:29:127

    ... which is roughly equal to the progression 1:4:17

    ... which is roughly equal to $4^0:4^1:4^2$, so $n = 4$

Bradford's law

- The mathematical proportion between the number of journals in the core and in zone 2 is a constant = n *(Bradford's multiplier)*; and to zone 2 n^2
- Bradford expressed this relationship as $\mathbf{1{:}n{:}n^2}$, but he never defined the law mathematically – n was set at approx. 5
- Simple formulation: an 80/20 rule, that is, 20% of the journals in the bibliography covers 80% of the articles in the bibliography on a given subject

Ex: Scattering of article on 'Digital Libraries' among 27 journals during period of 1991 - 2009

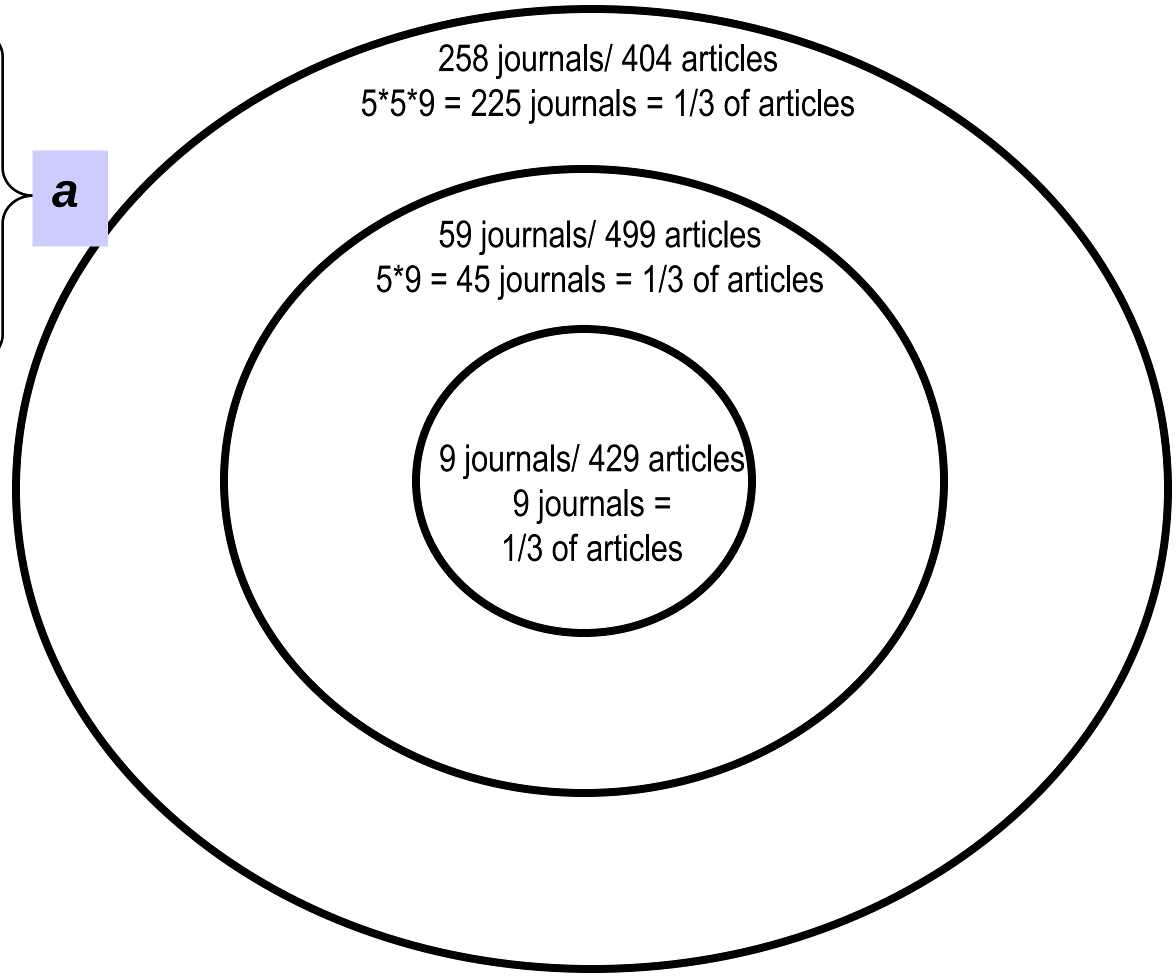
Sl. No.	Name of the Journals	No. of Articles	Percentage
1	The Electronic Library	66	14.90
2	Library Hi Tech New	52	11.74
3	Library Review	45	10.16
4	Online Information Review	44	9.93
5	Program: electronic library and information systems	42	9.48
6	OCLC Systems & Services	33	7.45
7	Library Management	27	6.09
8	New Library world	17	3.84
9	Journal of Documentation	17	3.84
10	Interlending & Document Supply	14	3.16
11	VINE	13	2.93
12	Reference Reviews	12	2.71
13	The Bottom Line: Managing Library Finances	11	2.48
14	Collection Building	9	2.03
15	Aslib Proceedings	7	1.58
16	Asian Libraries	7	1.58
17	Electronic Resources Review	6	1.35
18/	Reference Services Review	4	0.90
19	OCLC Systems & Services: International digital library perspectives	4	0.90
20	Performance Measurement and Metrics	3	0.68
21	Aslib Proceedings : New Information Perspectives	2	0.45
22	Journal of Knowledge Management	2	0.45
23	Internet Research	2	0.45
24	Kybernetes	1	0.23
25	Campus-Wide Information Systems	1	0.23
26	Digital libraries : Architecture, Information Modeling	1	0.23
27	Managing Digital Resources in Libraries	1	0.23
Total		**443**	**100.00**

Bradford's law

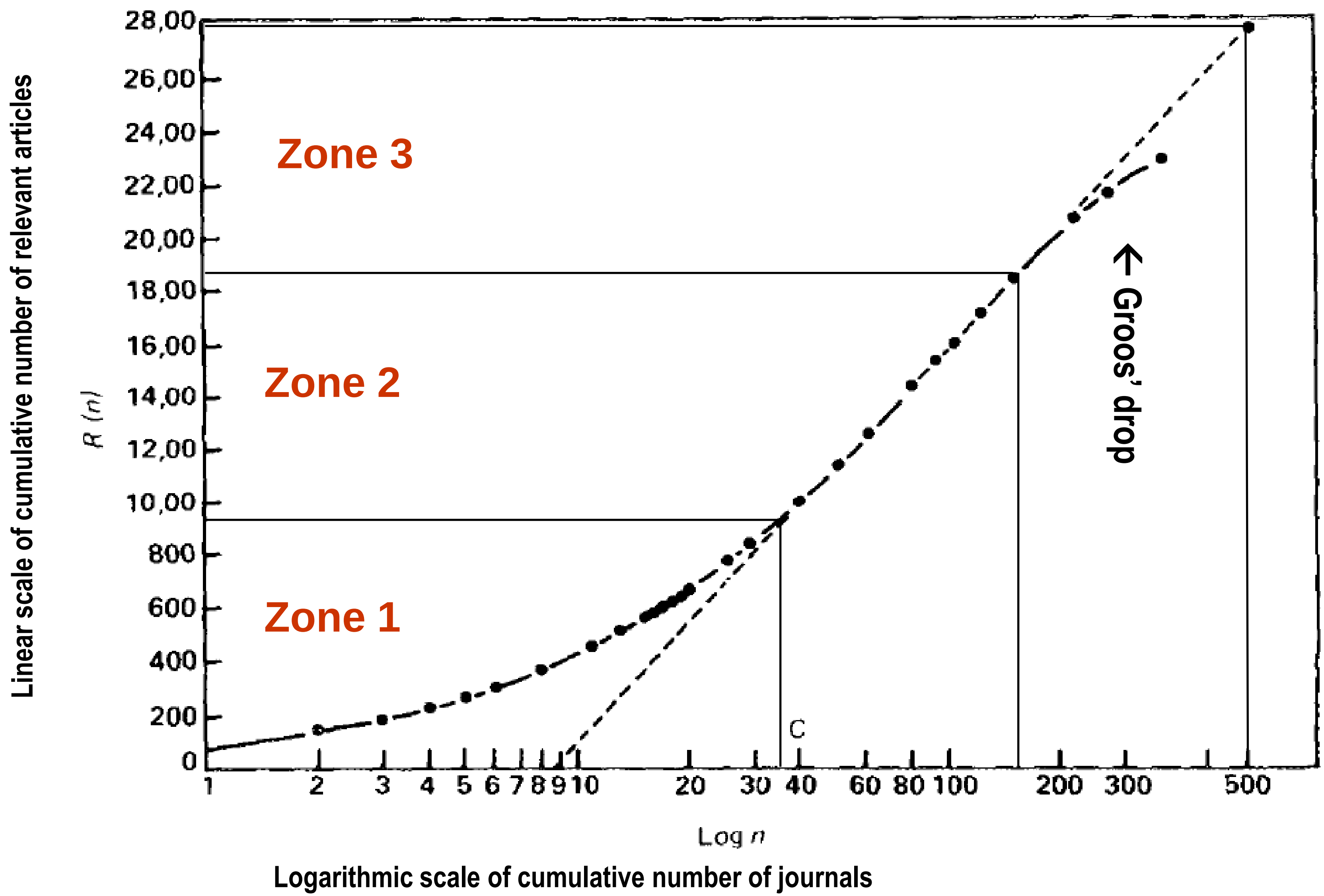
1	93	1	93
1	86	2	179
1	56	3	235
1	48	4	283
1	46	5	329
1	35	6	364
1	28	7	392
1	20	8	412
1	17	9	429
4	16	13	493
1	15	14	508
5	14	19	578
1	12	20	590
2	11	22	612
5	10	27	662
3	9	30	689
8	8	38	753
7	7	45	802
11	6	56	868
12	5	68	928
17	4	85	996
23	3	108	1065
49	2	157	1163
169	1	326	1332

a

$1, n, n^2 = 1, 5, 25$



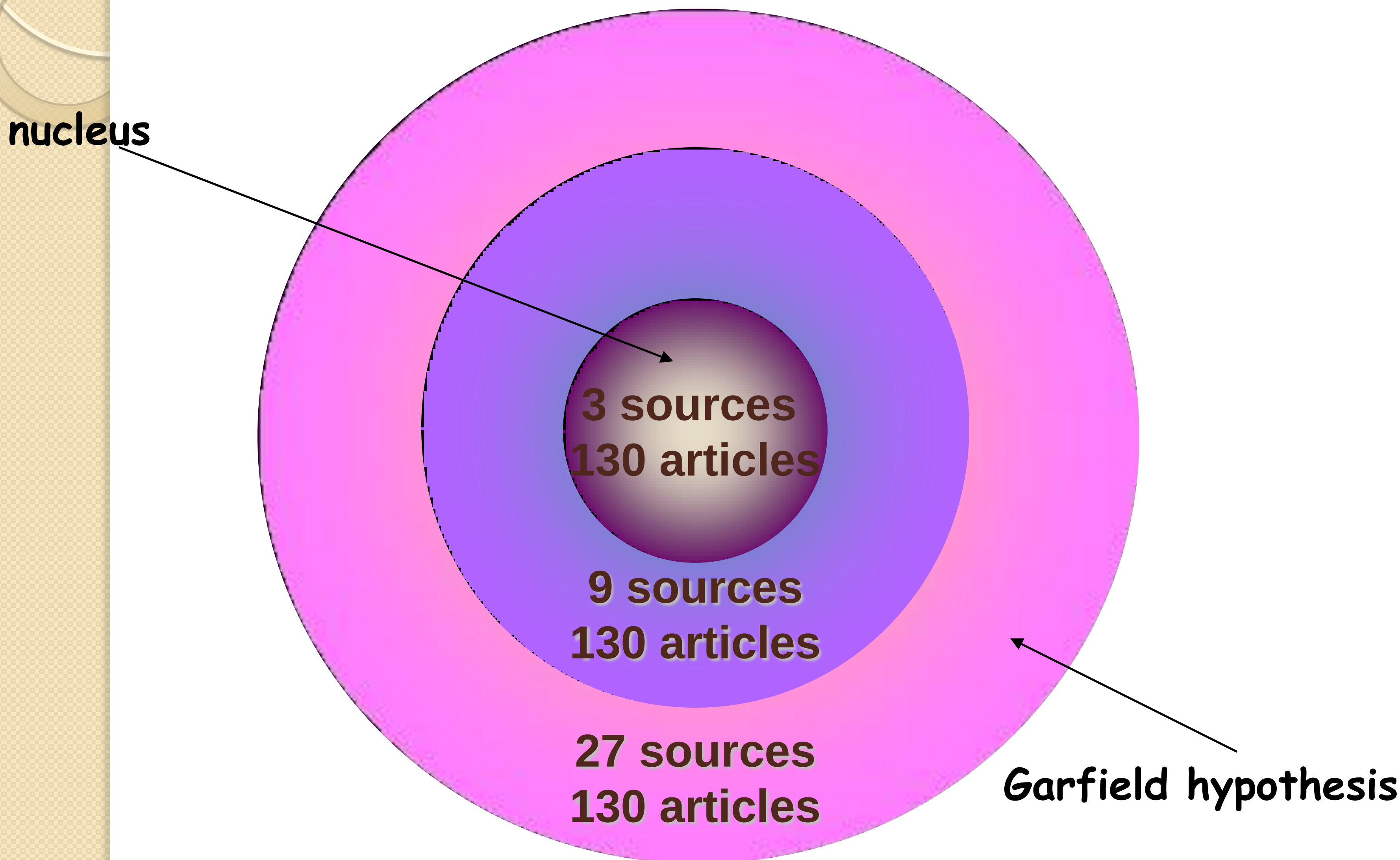
Bradford's law (graphical)



Bradford's Law of Scattering – an idealized example

No. of source journals		No. of articles per source	Total no. of articles	
3	1	60	60	130
	2	35	70	
9	1	30	30	130
	2	25	50	
	2	9	18	
	4	8	32	
27	10	6	60	130
	7	5	35	
	5	4	20	
	5	3	15	

Bradford's Law of Scattering – zones



Bradford's law

- Example: Bradford's geophysical data:

a is defined arbitrarily

	Zone 1	Zone 2	Zone 3
Articles	429	499	404
Journals	9 (=a)	59	258
	1	n	n^2
	1a	na	n^2a
	(1*9)	(5*9)	(25*9)

- The number of articles should ideally be the same in the three zones
- Bradford's law is not statistically exact, but is very widespread as a rule-of-thumb

Journal Wise Distributions of Publications

Sl. No.	Journal	Records	%	TLCS	TGCS	TLCR	Rank
1	Green Chemistry	355	3.96	1870	12340 (1)	869	1
2	Tetrahedron Letters	262	2.93	434	5206	348	2
3	Chemistry-A European Journal	162	1.81	266	6119(4)	174	3
4	Journal of Chemical Education	158	1.76	382	946	372	4
5	RSC Advances	146	1.63	34	731	266	5
6	Advanced Synthesis & Catalysis	133	1.48	237	3723	215	6
7	Angewandte Chemie-International Edition	128	1.43	357	8185(2)	103	7(1)
8	SYNLETT	122	1.36	136	1825	109	8
9	Inorganic Chemistry	113	1.26	41	3008	31	9
10	Chemsuschem	112	1.25	128	2095	169	10
11	European Journal of Organic Chemistry	109	1.22	105	1720	130	11
12	Tetrahedron	107	1.19	151	2234	215	12
13	Pure and Applied Chemistry	102	1.14	255	1916	162	13
14	Journal of the American Chemical Society	99	1.10	172	6261(3)	95	14(2)
15	Synthetic Communications	98	1.09	53	702	105	15
16	Journal of Industrial and Engineering Chemistry	70	0.78	17	518	39	16
17	Synthesis-Stuttgart	69	0.77	60	1010	79	17
18	Green Chemistry Letters and Reviews	64	0.71	34	291	127	18
19	Journal of Organic Chemistry	63	0.70	160	2314	73	19
20	Journal of Molecular Catalysis A-Chemical	61	0.68	94	1679	109	20
21	ACS Sustainable Chemistry & Engineering	60	0.67	21	275	139	21
22	Chemical Communications	59	0.66	141	2862	82	22
23	Current Organic Chemistry	59	0.66	37	389	165	22
24	European Journal of Inorganic Chemistry	59	0.66	17	641	37	22
25	Quimica Nova	59	0.66	118	320	205	22

Bradford's Law of scattering (1505Journals) of Food Chemistry Research Publications

S.No	No. of Journals	No. of Articles	Total no. of Articles	Cumulative No. of Articles
1	1	355	355	355
2	1	262	262	617
3	1	162	162	779
4	1	158	158	937
5	1	146	146	1083
6	1	133	133	1216
7	1	128	128	1344
8	1	122	122	1466
9	1	113	113	1579
10	1	112	112	1691
11	1	109	109	1800
12	1	107	107	1907
13	1	102	102	2009
14	1	99	99	2108
15	1	98	98	2206
16	1	70	70	2276
17	1	69	69	2345
18	1	64	64	2409
19	1	63	63	2472
20	1	61	61	2533
21	1	60	60	2593
22	4	59	236	2829
23	1	56	56	2885
24	1	51	51	2936
25	3(30)	50	150	3086(2984)
26	2	49	98	3184
27	2	48	96	3280
28	1	47	47	3327
29	1	46	46	3373
30	2	44	88	3461
31	2	43	86	3547
32	1	42	42	3589
33	2	40	80	3669
34	3	38	114	3783
35	2	36	72	3855

Cont…,

S.No	No. of Journals	No. of Articles	Total no. of Articles	Cumulative No. of Articles
36	2	35	70	3925
37	4	34	136	4061
38	1	33	33	4094
39	1	32	32	4126
40	4	31	124	4250
41	3	30	90	4340
42	2	29	58	4398
43	1	28	28	4426
44	4	27	108	4534
45	5	25	125	4659
46	3	24	72	4731
47	3	23	69	4800
48	6	22	132	4932
49	7	21	147	5079
50	3	20	60	5139
51	7	19	133	5272
52	6	18	108	5380
53	6	17	102	5482
54	8	16	128	5610
55	3	15	45	5655
56	6	14	84	5739
57	5	13	65	5804
58	7	12	84	5888
59	12(127)	11	132	6020(2934)
60	15	10	150	6170
61	18	9	162	6332
62	22	8	176	6508
63	19	7	133	6641
64	25	6	150	6791
65	45	5	225	7016
66	75	4	300	7316
67	129	3	387	7703
68	249	2	498	8201
69	751(1348)	1	751	8952(2932)
Total	1505	3846	8952	

Bradford's Distribution of Journals in Food Chemistry-Zone Analysis

Zone	No. of Journals	No. of Articles	Multiplier Factor
Z_1	30(1.99)	3086(34.47)	-
Z_2	127(8.44)	2934(32.77)	4.23
Z_3	1348(89.57)	2932(32.75)	10.61
Total	1505	8952	14.84(7.42)

George Kingsley Zipf 1935, 1949

- *The psycho-biology of language: an introduction to dynamic philology* (1935)
- *Human behavior and the principle of least effort: An introduction to human ecology* (1949)
- Looked, among others, at frequency distributions of words in given texts
  - counted distribution in James Joyces' *Ulysses*
- Provided an explanation as to why the found distributions happen:

  Principle of least effort

Zipf's law

- Zipf's law is often used to predict word frequencies in full text
- The law states that if you (in a relatively large text corpus):
  - "…list the words occurring within that text in order of decreasing frequency, the rank of a word on that list multiplied by its frequency will equal a constant."

(Zipf, 1949)

- ***r x f = c***, where ***r*** is the rank number and ***f*** the frequency of each term
- (happens to describe the linear part of a Bradford curve)
- Zipf's law is also not statistically exact, but it is extremely valuable for term weighting purposes in search engines with automatic indexing

Zipf's law: $r \cdot f = c$

- Where:

  r = rank (in terms of frequency)

  f = frequency (no. of times the given word is used in the text)

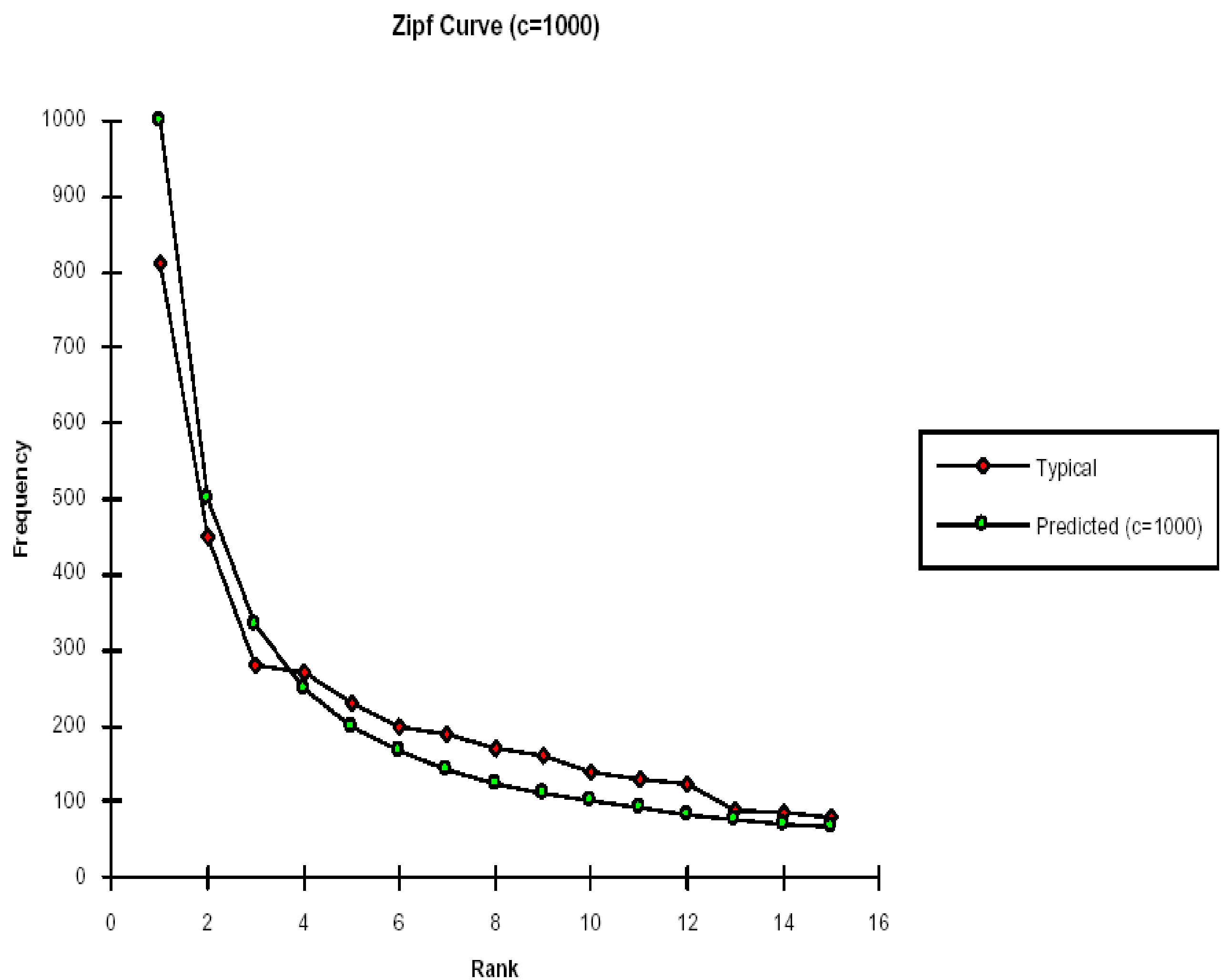
  c = constant for the given text

- For a given text the rank of a word multiplied by the frequency is a constant
- Works well for high frequency words, not so well for low – thus a number of modifications

Zipf's law

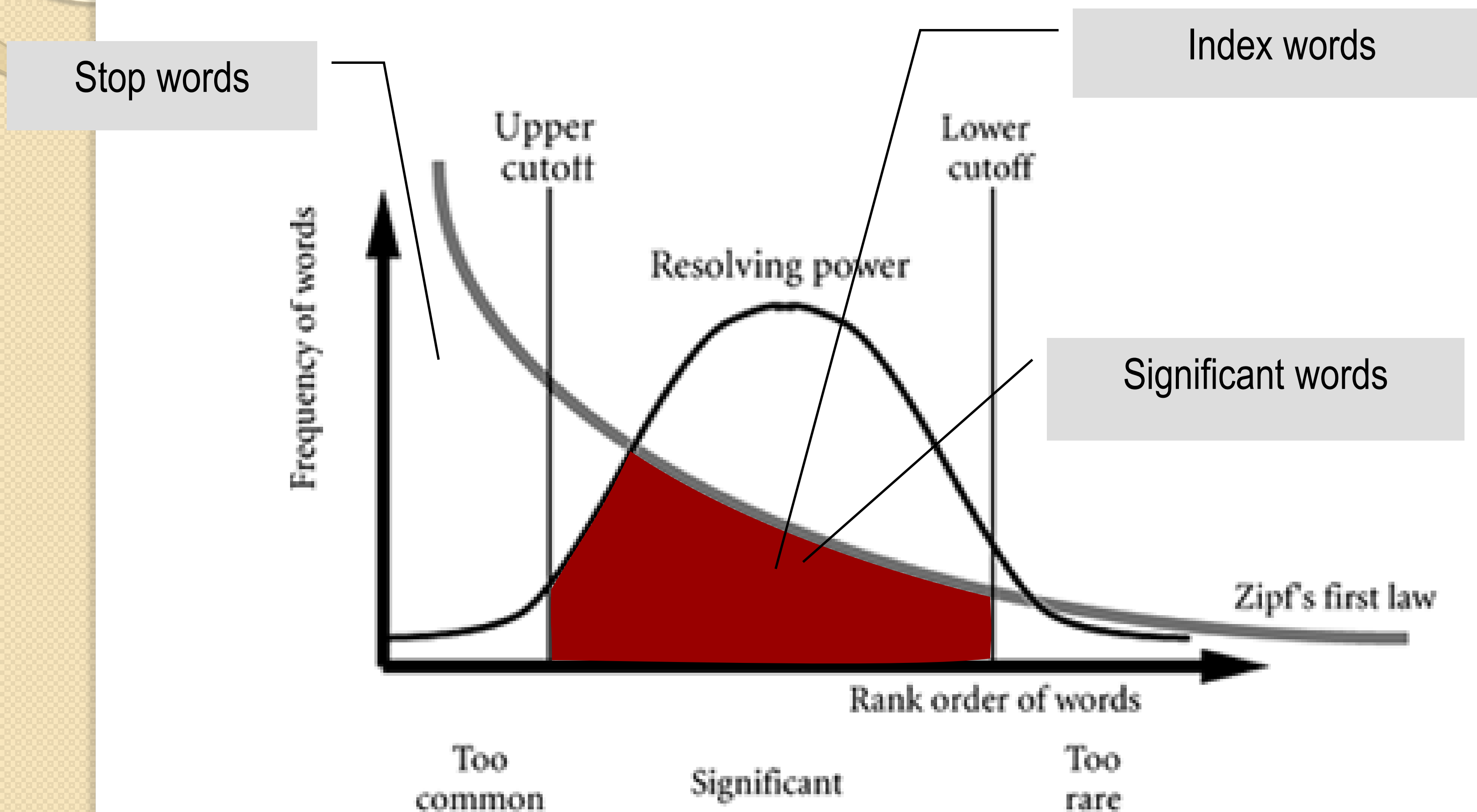
	B	C	D	E	F
2	Term	Rank	Typical	c=r*f	Predicted (c=1000)
3	*the*	1	810	810	1000
4	*of*	2	450	900	500
5	*a*	3	280	840	333
6	*information*	4	270	1080	250
7	*is*	5	230	1150	200
8	*to*	6	200	1200	167
9	*and*	7	190	1330	143
10	*that*	8	170	1360	125
11	*as*	9	160	1440	111
12	*in*	10	140	1400	100
13	*we*	11	130	1430	91
14	*be*	12	125	1500	83
15	*or*	13	90	1170	77
16	*may*	14	85	1190	71
17	*by*	15	80	1200	67

Zipf's law



Zipf Curve (c=1000)
1000
900
800
700
600
500
400
300
200
100
0
Frequency
0
2
4
6
8
10
12
14
16
Rank
Typical
Predicted (c=1000)

H.P. Luhn's *Resolving power & cut-off levels* for significant words (1957)



Charles F. Gosnell 1944 Obsolescence

- He studied obsolescence of books in academic libraries via their use
    - *College Res. Libr*. (1994) 5:115-125
- But this was extended to study of articles via citations, and other sources
- Age of citations in articles in a subject:
    - half life – half of the citations are x year old etc
        - different subjects have very different half-lives

References:

1. Almind, T.C. and Ingwersen, P. (1997). Informetric analysison the World Wide Web:methodological approachesto “webometrics”. Journal of Documentation, 53 (4), 404- 426.

2. Bar-Ilan, Judit. and Peritz, Bluma C. (2009.) A method for measuring the evolution of a topic on the web: The case of “informetrics”. *Journal of the American Society for Information Science and Technology, 60*(9), 1730–1740.

3. Bar-Ilan,Judit.(2008). Informetrics at the beginning of the 21stcentury: A review. *Journal of Informetrics, 2*(1), 1-52.

4. Björneborn , Lennart and Ingwersen, Peter. (2004). Toward a basic framework for webometrics.*Journal of the American Society for Information Science and Technology, 55*(14) 1216-1227.

5. Burrell,Quentin L. (2005). Measuring similarity of concentration between different informetric distributions: Two new approaches. *Journal of the American Society for Information Science and Technology, 56*(7), 704–714.

6. Egghe, L. (2005). Expansion of the field of informetrics: Origins and consequences. Information *Processing and Management, 41*(6), 1311-1316.

7. Egghe, Leo and Rousseau, Ronald. (1990). *Introduction to informetrics: Quantitative methods in library, documentation and information science*. Amsterdam: Elsevier Science Publishers.

Cont…

8. Erar, Aydn. (2002). Bibliometrics or informetrics: Displaying regularity in scientific patterns by using statistical distributions. *Hacettepe Journal of Mathematics and Statistics, 31,* 113- 125.

9. Jacobs, Daisy. (2010). Demystification of bibliometrics, scientometrics, Informetrics and webometrics. Paper presented at *11th DIS Annual Conference 2010, 2nd – 3rd September, Richardsbay, University of Zululand, South Africa*.

10. Nisonger, Thomas E. (2001). Report on the 8th international conference on scientometrics and informetrics in Sydney, Australia*. Library Collections, Acquisitions, and Technical Services, 25*(4), 353-494.

11. Rao, I.K. Ravichandra. (1998). Informetrics: Scope, definition, methodology and conceptual questions. Paper presented at *Workshop on Informetrics and Scientometrics 16-19 March 1998, Bangalore.*

12. Sangam, S.L et. al. (Eds.) (2010). *Webometrics, informetrics and scientometrics : Measuring scientific and technological progress of India*. Dharwad: Department of Library and Information Science, Karnataka University.

13. Stock,Wolfgang G. and Weber,Sonja. (2006). Facets of informetrics. *Information, 57*(8), 385-389.

14. Wilson, Concepcion s and Hood, William W. (2001). The literature of bibliometrics, scientometrics,and informetrics. *Scientometrics, 52*(2), 291-314.

15. Wormell, Irene. (1998).Informetrics: an emerging subdiscipline in information science. *Asian Libraries, 7*(10), 257-268.

THANK YOU