Lab 2: Analysis of the degree distribution

Ankit Tewari & Rodrigo Arias Mallo

October 18, 2017

1 Introduction

	Language	N	Max. deg.	M/N	N/M
1	Arabic	21065.00	2249.00	3.35	0.30
2	Basque	11868.00	576.00	2.18	0.46
3	Catalan	35524.00	5522.00	5.75	0.17
4	Chinese	35563.00	7645.00	5.20	0.19
5	Czech	66014.00	4727.00	3.97	0.25
6	English	29172.00	4547.00	6.86	0.15
7	Greek	12704.00	1081.00	3.52	0.28
8	Hungarian	34600.00	6540.00	3.10	0.32
9	Italian	13433.00	2678.00	4.23	0.24
10	Turkish	20403.00	6704.00	2.31	0.43

Table 1: Summary of the properties of the degree sequences. N is the number of nodes, M/N is mean degree where M is the sum of degrees.

	Language	λ	q	γ_1	γ_2	k_{max}
1	Arabic	3.22	0.30	2.10	2.05	49.83
2	Basque	1.83	0.46	2.36	2.33	48.88
3	Catalan	5.73	0.17	1.92	1.81	43.36
4	Chinese	5.17	0.19	1.89	1.76	43.29
5	Czech	3.89	0.25	2.05	1.98	49.35
6	English	6.85	0.15	1.80	1.64	43.08
7	Greek	3.41	0.28	2.13	2.08	49.80
8	Hungarian	2.93	0.32	2.34	2.30	48.93
9	Italian	4.16	0.24	2.11	2.05	49.83
10	Turkish	2.00	0.43	2.54	2.53	48.53

Table 2: Summary of the most likely parameters. γ_1 and γ_2 refer, respectively, to the exponent of the zeta distribution and the right-truncated distribution.

	Language	1	2	3	4	5
1	Arabic	240689.90	24557.59	543.93	371.55	0.00
2	Basque	50230.42	8429.94	911.24	66.06	0.00
3	Catalan	915591.88	63593.52	1937.69	1725.75	0.00
4	Chinese	620399.06	50729.48	2554.65	2065.35	0.00
5	Czech	942170.75	92592.43	1640.32	1501.88	0.00
6	English	742138.86	48299.57	4027.05	2602.43	0.00
7	Greek	157329.71	17326.31	356.58	195.90	0.00
8	Hungarian	468469.92	53687.35	2438.00	217.72	0.00
9	Italian	246034.50	23108.88	349.35	231.47	0.00
10	Turkish	193394.78	24665.10	2790.22	49.75	0.00

Table 3: The AIC difference (Δ) of a model on a given source. Models are numbered according to Table 2