

名前の分布

早川良¹, 福岡勇太¹, 廣澤航輝², 水口毅¹
 大阪府立大学工学部¹, 総合研究大学院大学高エネルギー加速器科学研究科²

人は誰でも名前—姓と名—を持っている。姓には多くの種類があり、鈴木や佐藤のように比較的ありふれたものもあれば、珍しいものもある。それらはどのような分布をしているのだろうか。宮島らが日本人の姓を対象に定量的な調査を行った結果、そのサイズ分布がべき則に従うということが明らかにされた¹⁾。その後、よりよいフィッティングを求める研究や、他の国や地域での分布に関する報告、べき則が生まれるメカニズムに対する研究などが行われている。

姓と同様名前にも様々な種類があり、ありふれたものから珍しいものまでである。では、そのサイズ分布はどうなっているのだろうか？我々はデータベースとして研究者リストと電話帳を用い、姓および名前のサイズ頻度分布を調査した²⁾。その結果、名前も姓と同様にその希少領域でべき的な分布を示すことが判明した(図1)。興味深いことに、その指数もよく似ている。

我々は、この名前のサイズ分布のべき的依存性が、数多くの命名過程が繰り返された結果生成された物であると考え、Yule過程と排他律をくみあわせた単純な数理モデルを提案した。まずGalton-Watson型の単性人口集団を考え、子供の数はポアソン分布であたえられるとする。新しく生まれた個体に(姓ではなく)名前を割り当てる。その際、まったく新しい名前を付ける場合と既存の名前から選ぶ場合とを確率的に選択する。既存の名前から選ぶ場合には、それまでに使われた回数に応じて選ばれる確率が変化すると仮定する。具体的には名前*i*が選ばれる確率*p_i*は、*i*のサイズを*s_i*として

$$p_i = \frac{s_i^\beta}{\sum_i s_i^\beta}$$

で与えられるとした。ここでβはパラメータである。さらに、同一家族内において同一名を禁止する(排他律)。以上のルールを採用し、数値計算を行ったところ、名前のサイズ頻度分布がべき的になることを再現した(図2)。

講演では、分布関数のパラメータ依存性や、姓と名前が同じ特徴を示すことに対する解釈、他の種類に名前についての解析³⁾も簡単に紹介する。

参考文献

- 1) S. Miyajima, et al., JPSJ **68** (1999) 3244-3247; S. Miyajima et al., et al., Physica **A278** (2000) 282-288.
- 2) R. Hayakawa, Y. Fukuoka, and T. Mizuguchi, JPSJ **81** (2012) 094001.
- 3) K. Hirose and T. Mizuguchi, submitted.

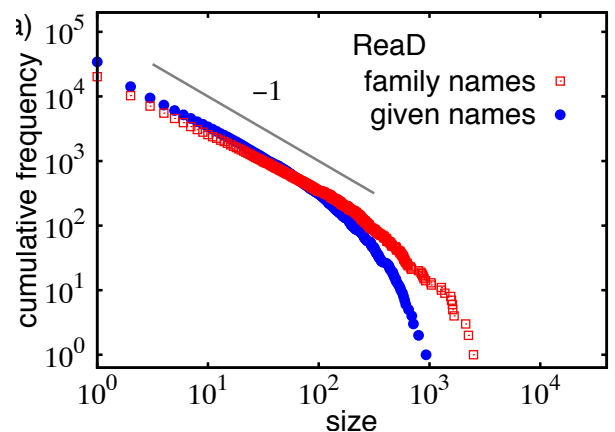


図1 日本人の名前のサイズ頻度分布。赤四角は名字。青丸は名前。横軸は名前のサイズで大きいほどありふれている。縦軸は累積分布。サイズの小さい領域で姓・名前いずれもべき的に振る舞う。

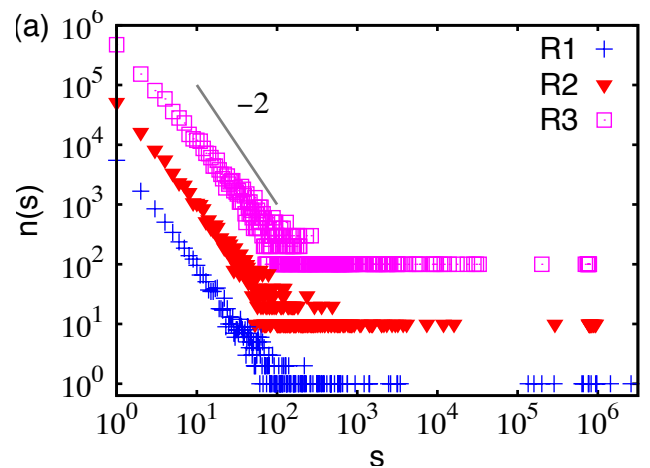


図2 モデルによる名前のサイズ分布。横軸はサイズで縦軸は分布関数でβ=1。R1は排他律なし。R2は親兄弟の名前まで排他的。R3は祖父叔父まで排他的。(論文2より転載)