

Distribution of Korean Names

Beom Jun Kim¹, Hye Jin Park¹, and Seung Ki Baek²

¹Department of Physics, Sungkyunkwan University, Suwon 440-746, Republic of Korea

²Department of Physics, Pukyong National University, Busan 608-737, Republic of Korea

E-mail: ¹beomjun@skku.edu

Abstract

Korean family name distribution is unique in the sense that the cumulative probability distribution is of the logarithmic form and the rank-frequency plot exhibits an exponential form. The origin of Korean distribution is studied within the master equation approach, and compared with empirical historical data for several hundred years. We also report our recent work on the time evolution of the popularity of Korean first names.

Keyword: Family names, First names, Distribution function, Power-law

Systems of family names have been developed in most countries in order to make distinctions among families for various biological, sociological, and economical reasons: For instances, marriages among close family members have disastrous genetic effects, most people want to keep the wealth of the family within the family, and giving a social identity (you are one of us) to family members has huge advantages in farming, construction, and forming a bigger army. Once introduced, the inheritance of the family name in most countries follows a very simple rule: children's family names are from the father. However, the time evolution of the family name distribution is not yet completely understood.

Korea exhibits exponentially decaying rank-size plots, or Zipf plots for the size versus the rank of families, and the number of family names increases logarithmically with population [1], whereas for most other countries where the empirical studies have been made, both follow a power-law form [2]. It is also shown that the Korean family name distribution was stable at least for several hundred years from the analysis of genealogical trees recorded in family books [3]. A recent study [2] has suggested the importance of the name generation rate in explaining the unique type of Korean family name distribution.

Very small number of Korea family names makes it impossible to distinguish one person from others only by using his/her family name. Accordingly, the homogeneity of Korean family names must be compensated by the diversity of first names. We investigate the first name distribution in Korea, and how the popularity of first names has changed for about a hundred years.

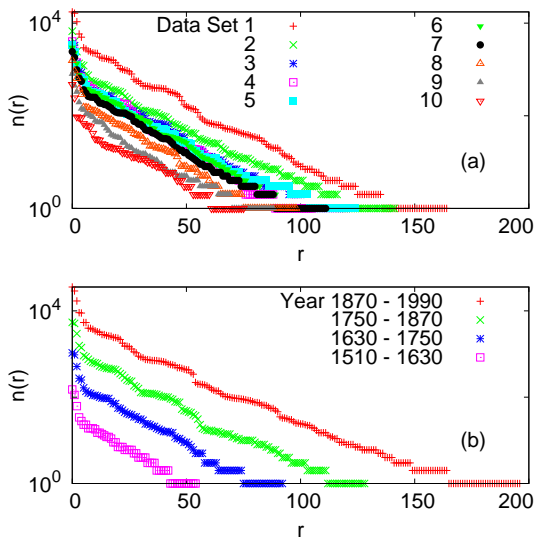


Figure 1: (a) The size $n(r)$ versus the rank r of families taken from the 10 family books. Each data set exhibits qualitatively the same exponentially decaying rank-size plot. (b) All data sets are merged and then split according to birth years. Overall feature has not changed for about five hundred years. (Figure from [3].)

References

- [1] B.J. Kim and S.M. Park, "Distribution of Korean Family Names", *Physica A* 347, 683 (2005).
- [2] S.K. Baek, H.A.T. Kiet, and B.J. Kim, "Family name distributions: Master equation approach", *Phys. Rev. E* 76, 046113 (2007).
- [3] H.A.T. Kiet, S.K. Baek, H. Jeong, and B.J. Kim, "Korean family name distribution in the past", *J. Korean Phys. Soc.* 51, 1812 (2007).