

MODELING THE CHANGING POPULARITY OF NAMES

by

Bengt Sigurd, Mats Eeg-Olofsson and Jørgen Ouren

The popularity of Norwegian first names 1880-2000 can be studied thanks to official Norwegian statistics. The most common curve shows a fast rise and slow fall, which can be approximated by a mathematical gamma function. The curve presumably reflects the parents' changing enthusiasm for a particular name.

Introduction

The Norwegian statistical agency (Statistisk Sentralbyrå) offers a data base on the Internet at www.ssb.no/emner/00/navn/, where the frequencies of several hundred Norwegian first names from 1880-2000 are shown in diagrams. As we have noted earlier, there is a certain recycling of names in Scandinavia and quite a few names, e.g. *Martin*, *Kristian* and *Lars* in Sweden, *Kristine* and *Karoline* in Norway, reappear after about 120 years (see publications by the Norwegian statistical agency; Sigurd and Eeg-Olofsson 2004). Also interesting is the shape of the historical frequency diagrams, as most of them display a fast frequency rise followed by a longer, slow fall. Such a shape can be approximated and modeled by a frequency function based on the mathematical gamma distribution. We will illustrate typical frequency curves for names, and show an approximating curve and gamma function which fits the name *Sverre* well. With somewhat different parameters, the function should fit several other names and allow us to predict the development of the popularity of a name.

The study of the developmental frequency patterns of names is interesting, since the same patterns are likely to show up in other fashion behavior. Modeling the patterns is not only of linguistic and sociological, but also of commercial interest, as it makes it possible to predict the development of a fashion or the success of a new product.

Types of curves

The name curves generally look like hills with a discernible rise and a fall (even though the whole 'hill' is not always visible in the Norwegian diagrams, which only include frequencies from 1880 to 2000, i.e. a span of a mere 120 years). In several cases, only the fall of a preceding wave or the rise of a second wave can be seen; in other cases, we can see both. See the diagram for *Sofie* (fig. 1).

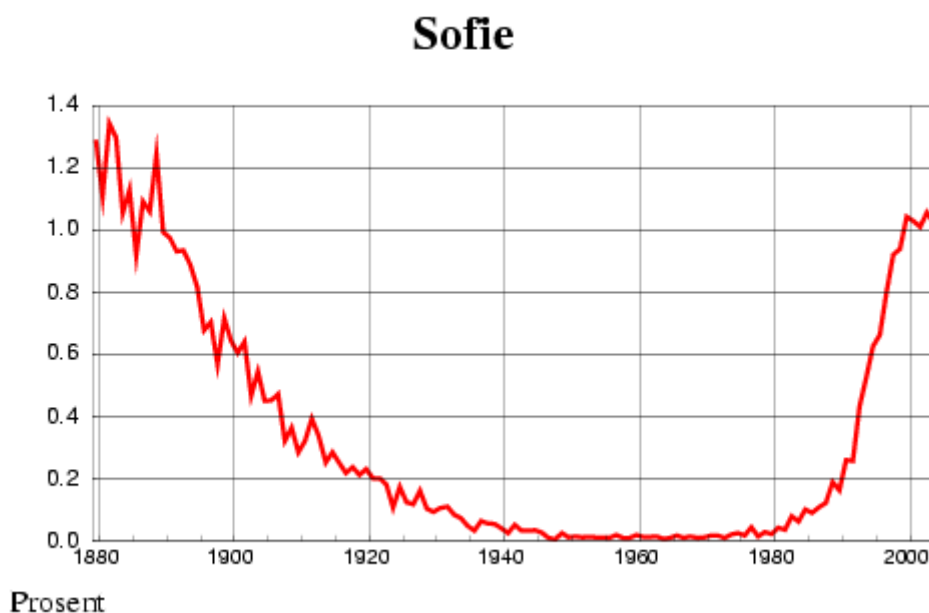


Fig. 1. Frequency of children given the name *Sofie* in different years. (Source: Statistisk Sentralbyrå).

It is natural to distinguish between symmetric and asymmetric frequency curves. Symmetric names show almost equally long rise and fall parts. This type is illustrated by the name *Åse*, which appeared around 1880 and rose to its top after 60 years, around 1940.

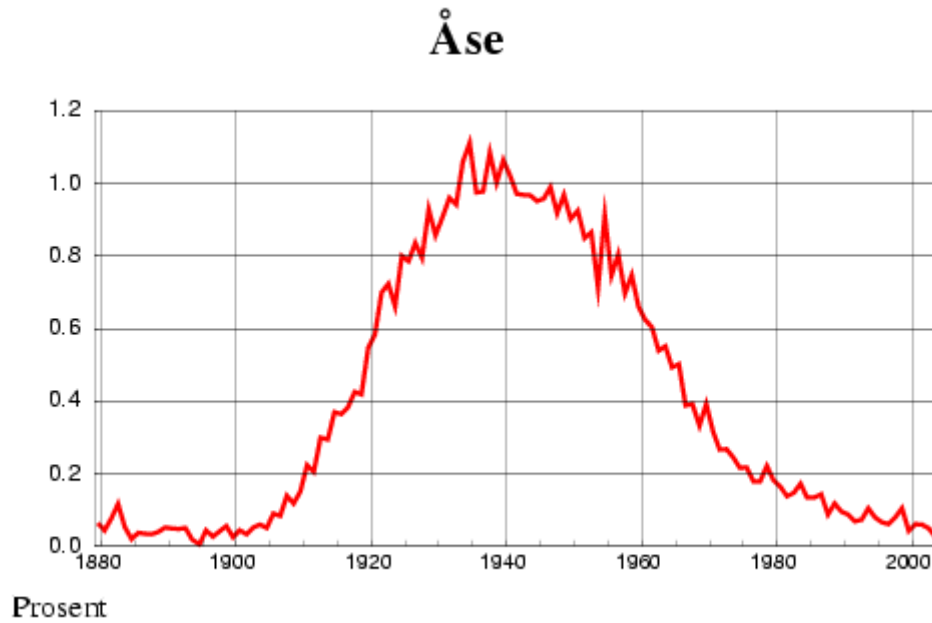


Fig. 2. Frequency of children given the name *Åse* in different years. (Source: Statistisk Sentralbyrå).

Asymmetric names may have a slow increase in the beginning, followed by a fast fall. This type is not common, but *Øystein* is an example.

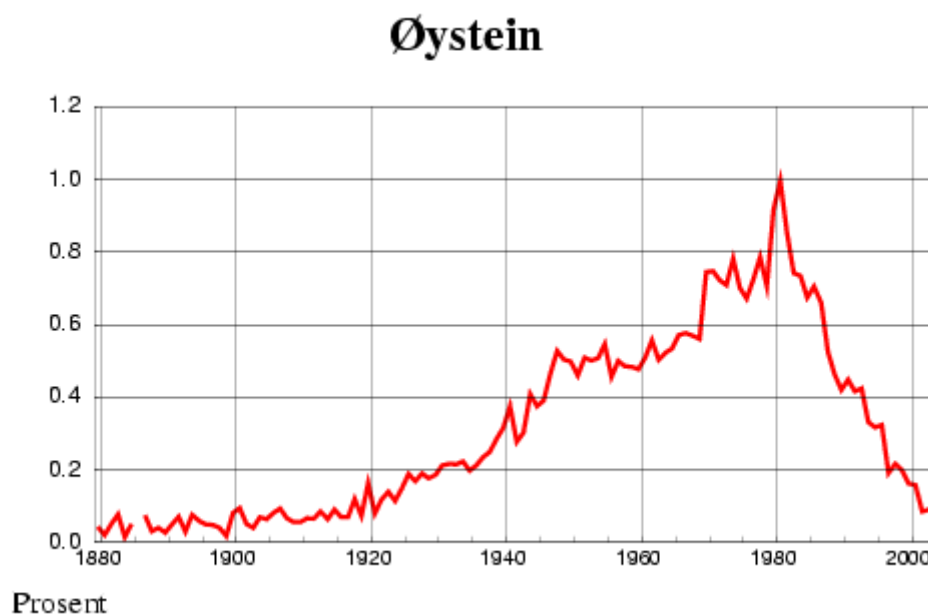


Fig. 3. Frequency of children given the name *Øystein* in different years. (Source: Statistisk Sentralbyrå).

The most common types of names, however, show a fast rise followed by a slow, gradual decrease. Out of the 150 male names displayed on the first page of the web site, about 100 can be classified as fast-slow names or 'gamma names' (as defined below). The same is true for the female names.

Gamma names

The frequency pattern of *Sverre* displays a development which reflects a beginning enthusiasm, rising fast to a top (i.e. after ca. 30 years). This is followed by a subsequent, gradual lack of interest, possibly as parents note that there are many children around with that name. The top percentage is generally 0.5-2.5. The decrease of the frequency looks like an exponential curve more or less approaching zero. The data and an approximation to a curve are shown in fig. 4.

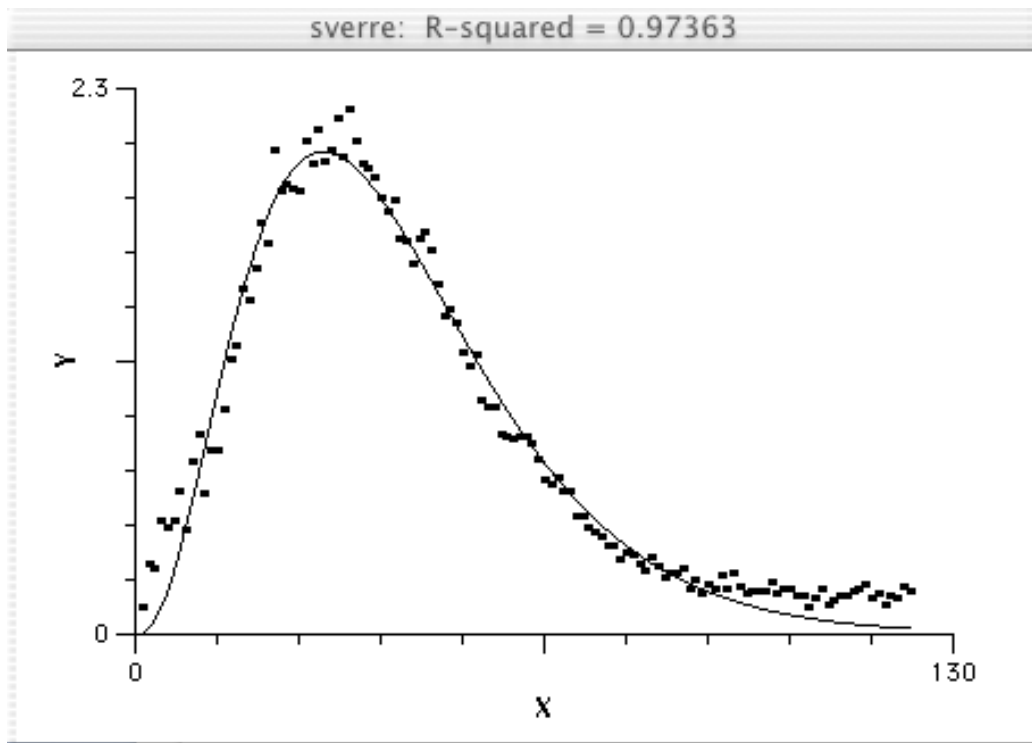


Fig. 4. Frequency of children given the name *Sverre* in different years, together with a curve approximation. (Source of data: Statistisk Sentralbyrå).

The curve for *Sverre* can be approximated by a frequency curve based on the mathematical gamma distribution, as shown in the diagram of fig. 4. (The gamma distribution is applied to a different kind of data in Sigurd, Eeg-Olofsson and van de Weijer (2004), where the general characteristics of the gamma distribution are discussed in more detail).

The function for the gamma distribution used here has the form:

$$f = a \cdot t^b \cdot c^t$$

where f is frequency, a is a scaling factor, t is time in years, b is a suitable parameter having to do with the early rise of the curve, and c is a parameter < 1 having to do with the exponential decrease of the frequency. The sign $^$ is used before an exponent and means 'raised to'.

The frequency rises with the increase of time, but this is counteracted by the latter part of the formula, in which the frequency decreases with time. For the name *Sverre*, the following function gives a fairly good fit. If we decrease the exponential value from 0.916608 to e.g. 0.90, the curve falls faster. The correlation of the function is as good as $r^2 = 0.97363$.

$$f = 0.0039568 * t^{2.60327} * 0.916608^t$$

Some of the data and approximated values for Sverre are shown in the following lists.

Year	Data	Approx
1881	0.118	0.004
1885	0.454	0.169
1890	0.840	0.664
1895	1.162	1.235
1900	1.730	1.690
1905	1.880	1.955
1910	1.989	2.033
1915	2.076	1.965
1920	1.781	1.800
1925	1.668	1.582
1930	1.366	1.347
1935	0.983	1.117
1940	0.820	0.906
1945	0.649	0.722
1950	0.501	0.567
1955	0.376	0.439
1960	0.300	0.336
1965	0.262	0.254
1970	0.177	0.191
1975	0.261	0.142
1980	0.186	0.105
1985	0.162	0.077
1990	0.129	0.056

MODELING THE CHANGING POPULARITY OF NAMES

1995	0.189	0.041
2000	0.163	0.030

Some further curves are shown combined in fig. 5. The curves for Jenny and Nina are of the gamma type, as one can see from the diagrams in fig. 5.

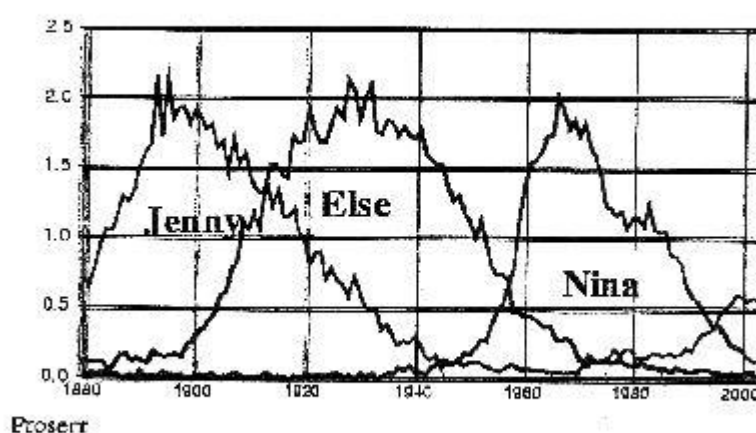


Fig. 5. Frequencies of *Jenny*, *Else*, *Nina*. (Source of data: Statistisk Sentralbyrå).

Discussion

The appearance of a name gaining popularity generally has to do with the existence of an idol: a royal person, a sports icon, a film star, an artist; here, we will not discuss which person(s) have started a particular name wave. Nor will we discuss why certain names show a sharp downfall. But it is clear that the popularity of the name *Adolf* fell by the end of World War II, and also, why the name *Harald* rose in popularity when the Norwegian Crown Prince (presently King Harald VI) was given that name upon his birth in 1937.

More data are needed in order to support the analysis of name frequencies in other cultures, but it seems highly reasonable to assume

that the gamma curve also is dominating in other cultures where names are given freely. Among further potential research tasks, we mention the study of the frequency development of fashion words and fads. New adjectives, especially those denoting 'good' or 'bad', are like fashions, and appear and fade away in similar ways. In Swedish, the expression *alla tiders* ('swell'; lit. 'of all times') was popular in the 1930s, but is completely outdated today (as is its US English counterpart). In similar ways, various positive adjectives have entered the English language during the 20th century, such as *swell*, *cool*, *sexy* etc.; it would be interesting to follow also their development.

*Department of Linguistics and Phonetics
Lund University
Helgonabacken 12
S-22362 Lund
Sweden*

References

- Sigurd, Bengt & Mats Eeg Olofsson. 2004. Modeväxlingar hos förnamn i Norden. *Arkiv för Nordisk Filologi* 119.209-225.
- Sigurd, Bengt, Mats Eeg-Olofsson & Joost van de Weijer. 2004. Word length, sentence length and frequency – Zipf revisited. *Studia Linguistica* 58(1).37-52.